Topology
Second Edition
by James Munkres

Solutions Manual

by Dan Whitman

April 14, 2019

Chapter 1 Set Theory and Logic

§1 Fundamental Concepts

Exercise 1.1

Check the distributive laws for $\cup$ and $\cap$ and DeMorgan's laws.

Solution:

Suppose that A, B, and C are sets. First we show that $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$.

Proof. We show this as a series of logical equivalences:

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

which of course shows the desired result. □

Next we show that $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.

Proof. We show this in the same way:

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

which of course shows the desired result. □

Now we show the first DeMorgan's law that $A - (B \cup C) = (A - B) \cap (A - C)$.

Proof. We show this in the same way:

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

which is the desired result. □

Lastly we show that $A - (B \cap C) = (A - B) \cup (A - C)$.

Proof. Again we use a sequence of logical equivalences:

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

as desired. □

Exercise 1.2

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" symbol 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) $A - (A - B) = B$.
(f) $A - (B - A) = A - B$.
(g) $A \cap (B - C) = (A \cap B) - (A \cap C)$.
(h) $A \cup (B - C) = (A \cup B) - (A \cup C)$.
(i) $(A \cap B) \cup (A - B) = A$.
(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) $A \times (B - C) = (A \times B) - (A \times C)$.
(p) $(A - B) \times (C - D) = (A \times C - B \times C) - A \times D$.
(q) $(A \times B) - (C \times D) = (A - C) \times (B - D)$.

Solution:

(a) We claim that $A \subset B$ and $A \subset C \Rightarrow A \subset (B \cup C)$ but that the converse is not generally true.

Proof. Suppose that $A \subset B$ and $A \subset C$ and consider any $x \in A$. Then clearly also $x \in B$ since $A \subset B$ so that $x \in B \cup C$. Since x was arbitrary, this shows that $A \subset (B \cup C)$ as desired.

To show that the converse is not true, suppose that $A = \{1, 2, 3\}$, $B = \{1, 2\}$, and $C = \{3, 4\}$. Then clearly $A \subset \{1, 2, 3, 4\} = B \cup C$ but it neither true that $A \subset B$ (since $3 \in A$ but $3 \notin B$) nor $A \subset C$ (since $1 \in A$ but $1 \notin C$). □

(b) We claim that $A \subset B$ or $A \subset C \Rightarrow A \subset (B \cup C)$ but that the converse is not generally true.

Proof. Suppose that $A \subset B$ or $A \subset C$ and consider any $x \in A$. If $A \subset B$ then clearly $x \in B$ so that $x \in B \cup C$. If $A \subset C$ then clearly $x \in C$ so that again $x \in B \cup C$. Since x was arbitrary, this shows that $A \subset (B \cup C)$ as desired.

The counterexample that disproves the converse of part (a), also serves as a counterexample to the converse here. Again this is because $A \subset B \cup C$ but neither $A \subset B$ nor $A \subset C$, which is to say that $A \not\subset B$ and $A \not\subset C$. Hence it is not true that $A \subset B$ or $A \subset C$. □

(c) We claim that this biconditional is true.

Proof. ($\Rightarrow$) Suppose that $A \subset B$ and $A \subset C$ and consider any $x \in A$. Then clearly also $x \in B$ and $x \in C$ since both $A \subset B$ and $A \subset C$. Hence $x \in B \cap C$, which proves that $A \subset B \cap C$ since x was arbitrary.

($\Leftarrow$) Now suppose that $A \subset B \cap C$ and consider any $x \in A$. Then $x \in B \cap C$ as well so that $x \in B$ and $x \in C$. Since x was an arbitrary element of A, this of course shows that both $A \subset B$ and $A \subset C$ as desired. $\square$

(d) We claim that only the converse is true here.

Proof. To show the converse, suppose that $A \subset B \cap C$. It was shown in part (c) that this implies that both $A \subset B$ and $A \subset C$. Thus it is clearly true that $A \subset B$ or $A \subset C$.

As a counterexample to the forward implication, let $A = \{1\}$, $B = \{1, 2\}$, and $C = \{3, 4\}$ so that clearly $A \subset B$ and hence $A \subset B$ or $A \subset C$ is true. However we have that B and C are disjoint so that $B \cap C = \varnothing$, therefore $A \not\subset \varnothing = B \cap C$ since $A \neq \varnothing$. $\square$

(e) We claim that $A - (A - B) \subset B$ but that the other direction is not generally true.

Proof. First consider any $x \in A - (A - B)$ so that $x \in A$ but $x \notin A - B$. Hence it is not true that $x \in A$ and $x \notin B$. So it must be that $x \notin A$ or $x \in B$. However, since we know that $x \in A$, it has to be that $x \in B$. Thus $A - (A - B) \subset B$ since x was arbitrary.

Now let $A = \{1, 2\}$ and $B = \{2, 3\}$. Then we clearly have $A - B = \{1\}$, and thus $A - (A - B) = \{2\}$. So clearly B is not a subset of $A - (A - B)$ since $3 \in B$ but $3 \notin A - (A - B)$. $\square$

(f) Here we claim that $A - (B - A) \supset A - B$ but that the other direction is not generally true.

Proof. First suppose that $x \in A - B$ so that $x \in A$ but $x \notin B$. Then it is certainly true that $x \notin B$ or $x \in A$ so that, by logical equivalence, it is not true that $x \in B$ and $x \notin A$. That is, it is not true that $x \in B - A$, which is to say that $x \notin B - A$. Since also $x \in A$, it follows that $x \in A - (B - A)$, which shows the desired result since x was arbitrary.

To show that the other direction does not hold consider the counterexample $A = \{1, 2\}$ and $B = \{2, 3\}$. Then $B - A = \{3\}$ so that $A - (B - A) = \{1, 2\} = A$. We also have that $A - B = \{1\}$ so that $2 \in A - (B - A)$ but $2 \notin A - B$. This suffices to show that $A - (B - A) \not\subset A - B$. $\square$

(g) We claim that equality holds here, i.e. that $A \cap (B - C) = (A \cap B) - (A \cap C)$.

Proof. ($\subset$) Suppose that $x \in A \cap (B - C)$ so that $x \in A$ and $x \in B - C$. Thus $x \in B$ but $x \notin C$. Since both $x \in A$ and $x \in B$ we have that $x \in A \cap B$. Also since $x \notin C$ it clearly must be that $x \notin A \cap C$. Hence $x \in (A \cap B) - (A \cap C)$, which shows the forward direction since x was arbitrary.

($\supset$) Now suppose that $x \in (A \cap B) - (A \cap C)$. Hence $x \in A \cap B$ but $x \notin A \cap C$. From the former of these we have that $x \in A$ and $x \in B$, and from the latter it follows that either $x \notin A$ or $x \notin C$. Since we know that $x \in A$, it must therefore be that $x \notin C$. Hence $x \in B - C$ since $x \in B$ but $x \notin C$. Since also $x \in A$ we have that $x \in A \cap (B - C)$, which shows the desired result since x was arbitrary. $\square$

(h) Here we claim that $A \cup (B - C) \supset (A \cup B) - (A \cup C)$ but that the forward direction is not generally true.

Proof. First consider any $x \in (A \cup B) - (A \cup C)$ so that $x \in A \cup B$ and $x \notin A \cup C$. From the latter, it follows that $x \notin A$ and $x \notin C$ since otherwise we would have $x \in A \cup C$. From the former, we have that $x \in A$ or $x \in B$ so that it must be that $x \in B$ since $x \notin A$. Therefore we have that $x \in B$ and $x \notin C$ so that $x \in B - C$. From this it obviously follows that $x \in A \cup (B - C)$, which shows that $A \cup (B - C) \supset (A \cup B) - (A \cup C)$ since x was arbitrary.

To show that the forward direction does not always hold, consider the sets $A = \{1, 2\}$, $B = \{2, 3\}$, and $C = \{2\}$. Then we clearly have that $B - C = \{3\}$, and hence $A \cup (B - C) = \{1, 2, 3\}$. On the other hand, we have $A \cup B = \{1, 2, 3\}$ and $A \cup C = \{1, 2\}$ so that $(A \cup B) - (A \cup C) = \{3\}$. Hence, for example, $1 \in A \cup (B - C)$ but $1 \notin (A \cup B) - (A \cup C)$, which suffices to show that $A \cup (B - C) \not\subset (A \cup B) - (A \cup C)$ as desired. □

(i) We claim that equality holds here.

Proof. We show this with a chain of logical equivalences:

$$\begin{aligned} x \in (A \cap B) \cup (A - B) &\Leftrightarrow x \in A \cap B \vee x \in A - B \\ &\Leftrightarrow (x \in A \wedge x \in B) \vee (x \in A \wedge x \notin B) \\ &\Leftrightarrow x \in A \wedge (x \in B \vee x \notin B) \\ &\Leftrightarrow x \in A \wedge \text{True} \\ &\Leftrightarrow x \in A\,, \end{aligned}$$

where we note that "True" denotes the fact that $x \in B \vee x \notin B$ is always true by the excluded middle property of logic. □

(j) We claim that this implication is true.

Proof. Suppose that $A \subset C$ and $B \subset D$. Consider any $(x, y) \in A \times B$ so that $x \in A$ and $y \in B$ by the definition of the cartesian product. Then also clearly $x \in C$ and $y \in D$ since $A \subset C$ and $B \subset D$. Hence $(x, y) \in C \times D$, which shows the result since the ordered pair (x, y) was arbitrary. □

(k) We claim that the converse of (j) is not always true.

Proof. Consider the following sets:

$$\begin{aligned} A &= \varnothing & \qquad C &= \{1\} \\ B &= \{1, 2\} & \qquad D &= \{2\}\ . \end{aligned}$$

Then we have that $A \times B = \varnothing$ since there are no ordered pairs (x, y) such that $x \in A$ (since $A = \varnothing$). Hence it is vacuously true that $(A \times B) \subset (C \times D)$. However, clearly it is not the case that $B \subset D$, and so, even though $A \subset C$, it is not true that $A \subset C$ and $B \subset D$. □

(l) We claim that the converse of (j) is true with the stipulation that A and B are both nonempty.

Proof. Suppose that $(A \times B) \subset (C \times D)$. First consider any $x \in A$. Then, since $B \neq \varnothing$, there is a $b \in B$. Then $(x, b) \in A \times B$ so that clearly also $(x, b) \in C \times D$. Hence $x \in C$ so that $A \subset C$ since x was arbitrary. An analogous argument shows that $B \subset D$ since A is nonempty. Hence it is true that $A \subset C$ and $B \subset D$ as desired. □

(m) Here we claim that $(A \times B) \cup (C \times D) \subset (A \cup C) \times (B \cup D)$ but that the other direction is not always true.

Proof. First consider any $(x, y) \in (A \times B) \cup (C \times D)$ so that either $(x, y) \in A \times B$ or $(x, y) \in C \times D$. In the first case $x \in A$ and $y \in B$ so that clearly $x \in A \cup C$ and $y \in B \cup D$. Hence $(x, y) \in (A \cup C) \times (B \cup D)$. In the second case we have $x \in C$ and $y \in D$ so that again $x \in A \cup C$ and $y \in B \cup D$ are still both true. Hence of course $(x, y) \in (A \cup C) \times (B \cup D)$ here also. This shows the result in either case since (x, y) was an arbitrary ordered pair.

To show that the other direction does not always hold, consider $A = B = \{1\}$ and $C = D = \{2\}$. Then we clearly have $A \times B = \{(1, 1)\}$ and $C \times D = \{(2, 2)\}$ so that $(A \times B) \cup (C \times D) = \{(1, 1), (2, 2)\}$. We also have $A \cup C = B \cup D = \{1, 2\}$ so that $(A \cup C) \times (B \cup D) = \{(1, 1), (1, 2), (2, 1), (2, 2)\}$. This clearly shows that $(A \times B) \cup (C \times D) \not\supset (A \cup C) \times (B \cup D)$ as desired. □

(n) We claim that the equality holds here.

Proof. We can show this by a series of logical equivalences:

$$\begin{aligned}
(x, y) \in (A \times B) \cap (C \times D) &\Leftrightarrow (x, y) \in A \times B \wedge (x, y) \in C \times D \\
&\Leftrightarrow (x \in A \wedge y \in B) \wedge (x \in C \wedge y \in D) \\
&\Leftrightarrow (x \in A \wedge x \in C) \wedge (y \in B \wedge y \in D) \\
&\Leftrightarrow x \in A \cap C \wedge y \in B \cap D \\
&\Leftrightarrow (x, y) \in (A \cap C) \times (B \cap D)
\end{aligned}$$

as desired. □

(o) We claim that equivalence holds here as well.

Proof. ($\subset$) First consider any $(x, y) \in A \times (B - C)$ so that $x \in A$ and $y \in B - C$. From the latter of these we have that $y \in B$ but $y \notin C$. We clearly then have that $(x, y) \in A \times B$ since $x \in A$ and $y \in B$. It also has to be that $(x, y) \notin A \times C$ since $y \notin C$ even though it is true that $x\ inA$. Therefore $(x, y) \in (A \times B) - (A \times C)$ as desired.

($\supset$) Now suppose that $(x, y) \in (A \times B) - (A \times C)$ so that $(x, y) \in A \times B$ but $(x, y) \notin (A \times C)$. From the former we have that $x \in A$ and $y \in B$. It then must be that $y \notin C$ since $(x, y) \notin (A \times C)$ but we know that $x \in A$. Then we have $y \in B - C$ since $y \in B$ but $y \notin C$. Since also $x \in A$, it follows that $(x, y) \in A \times (B - C)$ as desired. □

(p) We claim the equivalence hold for this statement.

Proof. ($\subset$) Suppose that $(x, y) \in (A - B) \times (C - D)$ so that $x \in A - B$ and $y \in C - D$. Then we have that $x \in A$, $x \notin B$, $y \in C$, and $y \notin D$. So first, clearly $(x, y) \in A \times C$. Then, since $x \notin B$, we have that $(x, y) \notin B \times C$, and hence $(x, y) \in A \times C - B \times C$. Since $y \notin D$, we also have that $(x, y) \notin A \times D$, and thus $(x, y) \in (A \times C - B \times C) - A \times D$. This clearly shows the desired result since (x, y) was arbitrary.

($\supset$) Now suppose that $(x, y) \in (A \times C - B \times C) - A \times D$ so that $(x, y) \in A \times C - B \times C$ but $(x, y) \notin A \times D$. From the former we have that $(x, y) \in A \times C$ and $(x, y) \notin B \times C$. Thus $x \in A$ and $y \in C$ so that it has to be that $x \notin B$ since $(x, y) \notin B \times C$ but we know that $y \in C$. It also must be that $y \notin D$ since $(x, y) \notin A \times D$ but $x \in A$. Therefore we have that $x \in A$, $x \notin B$, $y \in C$, and $y \notin D$, from which it readily follows that $x \in A - B$ and $y \in C - D$. Thus clearly $(x, y) \in (A - B) \times (C - D)$, which shows the desired result since (x, y) was arbitrary. □

(q) Here we claim that $(A \times B) - (C \times D) \supset (A - C) \times (B - D)$ but that the forward direction is not true in general.

Proof. First consider any $(x, y) \in (A - C) \times (B - D)$ so that $x \in A - C$ and $y \in B - D$. Thus we have $x \in A$, $x \notin C$, $y \in B$, and $y \notin D$. From this clearly $(x, y) \in A \times B$ but $(x, y) \notin C \times D$. Hence $(x, y) \in (A \times B) - (C \times D)$, which clearly shows the desired result since (x, y) was arbitrary.

To show that the forward direction does not hold, consider $A = \{1, 2\}$, $B = \{a, b\}$, $C = \{2, 3\}$, and $D = \{b, c\}$. We then clearly have the following sets:

$$\begin{aligned} A \times B &= \{(1, a), (1, b), (2, a), (2, b)\} & A - C &= \{1\} \\ C \times D &= \{(2, b), (2, c), (3, b), (3, c)\} & B - D &= \{a\} \\ (A \times B) - (C \times D) &= \{(1, a), (1, b), (2, a)\} & (A - C) \times (B - D) &= \{(1, a)\}\ . \end{aligned}$$

This clearly shows that $(A \times B) - (C \times D)$ is not a subset of $(A - C) \times (B - D)$. $\square$

Exercise 1.3

(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 for the statement "If $x > 0$, then $x^2 - x > 0$."

Solution:

(a) First we claim that the original statement is true.

Proof. Since $x < 0$ we clearly have that $x - 1 < x < 0$ as well. Then, since the product of two negative numbers is positive, we have that $x^2 - x = x(x - 1) > 0$ as desired. $\square$

The contrapositive of this is, "If $x^2 - x \leq 0$, then $x \geq 0$." This is of course also true by virtue of the fact that the contrapositive is logically equivalent to the original implication.

Lastly, the converse of this statement is, "If $x^2 - x > 0$, then $x < 0$." We claim that this is not generally true.

Proof. A simple counterexample of $x = 2$ shows this. We have $x^2 - x = 2^2 - 2 = 4 - 2 = 2 > 0$, but also clearly $x = 2 > 0$ as well so that $x < 0$ is clearly false. $\square$

(b) First we claim that this statement is false.

Proof. As a counterexample, let $x = 1/2$. Then clearly $x > 0$, but we also have $x^2 - x = (1/2)^2 - 1/2 = 1/4 - 1/2 = -1/4 < 0$ so that $x^2 - x > 0$ is obviously not true. $\square$

The contrapositive is then "If $x^2 - x \leq 0$, then $x \leq 0$," which is false since it is logically equivalent to the original statement.

The converse is "If $x^2 - x > 0$, then $x > 0$," which we claim is false.

Proof. As a counterexample, consider $x = -1$ so that $x^2 - x = (-1)^2 - (-1) = 1 + 1 = 2 > 0$. However, we also clearly have $x = -1 < 0$ so that $x > 0$ is not true. $\square$

Exercise 1.4

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:

These are all basic logical negations using existential quantifiers:

(a) There is an $a \in A$ where $a^2 \notin B$.

(b) For every $a \in A$, $a^2 \notin B$.

(c) There is an $a \in A$ where $a^2 \in B$.

(d) For every $a \notin A$, $a^2 \notin B$.

Exercise 1.5

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

Solution:

(a) The statement on the right is the definition of the statement on the left so of course the implication and its converse are true.

(b) The implication is generally false.

Proof. As a counterexample, consider $\mathcal{A} = \{\{1\}, \{2\}\}$. Then clearly $\bigcup_{A \in \mathcal{A}} A = \{1, 2\}$ so that $1 \in \bigcup_{A \in \mathcal{A}} A$, but 1 is not in A for every $A \in \mathcal{A}$ since $1 \notin \{2\}$. $\square$

However, the converse *is* true.

Proof. Suppose that $x \in A$ for every $A \in \mathcal{A}$. Since $\mathcal{A}$ is nonempty there is an $A_0 \in \mathcal{A}$. Then $x \in A_0$ since $A_0 \in \mathcal{A}$. Hence by definition $x \in \bigcup_{A \in \mathcal{A}} A$ since $x \in A_0$ and $A_0 \in \mathcal{A}$. $\square$

(c) The implication here is true.

Proof. Suppose that $x \in \bigcap_{A \in \mathcal{A}} A$ so that by definition $x \in A$ for every $A \in \mathcal{A}$. Since $\mathcal{A}$ is nonempty there is an $A_0 \in \mathcal{A}$ so that in particular $x \in A_0$. This shows the desired result since $A_0 \in \mathcal{A}$. $\square$

The converse is not generally true.

Proof. As a counterexample consider $\mathcal{A} = \{\{1, 2\}, \{2, 3\}\}$. Then $1 \in \{1, 2\}$ and $\{1, 2\} \in \mathcal{A}$, but $1 \notin \bigcap_{A \in \mathcal{A}} A$ since clearly $\bigcap_{A \in \mathcal{A}} A = \{2\}$. $\square$

(d) The statement on the right is the definition of the statement on the left so of course the implication and its converse are true.

Exercise 1.6

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

Solution:

Again these involve simple logical negations of both sides of the implications:

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

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

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

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

Exercise 1.7

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 $-$.

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

Solution:

First, we obviously have

$$\begin{aligned}
D &= A \cap (B \cup C) \\
E &= (A \cap B) \cup C \,,
\end{aligned}$$

noting that $D \neq E$ generally though they appear similar. Regarding F we have the following sequence of logical equivalences:

$$\begin{aligned}
x \in F &\Leftrightarrow x \in A \wedge (x \in B \Rightarrow x \in C) \\
&\Leftrightarrow x \in A \wedge (x \notin B \vee x \in C) \\
&\Leftrightarrow x \in A \wedge \neg(x \in B \wedge x \notin C) \\
&\Leftrightarrow x \in A \wedge \neg(x \in B - C) \\
&\Leftrightarrow x \in A \wedge x \notin B - C \\
&\Leftrightarrow x \in A - (B - C)
\end{aligned}$$

so that of course $F = A - (B - C)$.

Exercise 1.8

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:

We claim that if a finite set has n elements, then its power set has 2^n elements, which is why it is called the power set.

Proof. We show this by induction on the size of the set. For the base case start with the the empty set in which $n = 0$. Clearly the only subset of $\varnothing$ is the trivial subset $\varnothing$ itself so that $\mathcal{P}(\varnothing) = \{\varnothing\}$. This has $1 = 2^0 = 2^n$ element obviously, which shows the base case. Now suppose that the power set of any set with n elements has 2^n elements. Let A be a set with $n+1$ elements, noting that this is nonempty since $n+1 \geq 1$ since $n \geq 0$. Hence there is an $x \in A$. For any subset $B \subset A$, either $x \notin B$ or $x \in B$. In the first case B is a subset of $A - \{x\}$ and in the latter $B = \{x\} \cup C$ for some $C \subset A - \{x\}$. Therefore $\mathcal{P}(A)$ has twice the number of elements of $\mathcal{P}(A - \{x\})$, one half being just the elements of $A - \{x\}$ and the other being those elements with x added in. But $A - \{x\}$ has n elements since A has $n+1$, and hence $\mathcal{P}(A - \{x\})$ has 2^n elements by the induction hypothesis. Thus $\mathcal{P}(A)$ has $2 \cdot 2^n = 2^{n+1}$ elements, which completes the induction. $\square$

Using this, we can answer all of the specific questions. If a set has two elements, than its power set has $2^2 = 4$ elements. If it has one element, then its power set has $2^1 = 2$ elements, namely $\mathcal{P}(\{x\}) = \{\varnothing, \{x\}\}$. If a set has three elements then its power set has $2^3 = 8$ elements. Lastly, if a set has no elements (i.e. it is the empty set), then its power set has $2^0 = 1$ elements. As noted in the proof we have $\mathcal{P}(\varnothing) = \{\varnothing\}$.

Exercise 1.9

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

Solution:

In following suppose that A is a set and $\mathcal{B}$ is a nonempty collection of sets. For arbitrary unions, we claim that

$$A - \bigcup_{B \in \mathcal{B}} B = \bigcap_{B \in \mathcal{B}} (A - B).$$

Proof. The simplest way to show this is with a series of logically equivalent statements. For any x we have that

$$\begin{aligned}
x \in A - \bigcup_{B \in \mathcal{B}} B &\Leftrightarrow x \in A \wedge x \notin \bigcup_{B \in \mathcal{B}} B \\
&\Leftrightarrow x \in A \wedge \neg \exists B \in \mathcal{B}(x \in B) \\
&\Leftrightarrow x \in A \wedge \forall B \in \mathcal{B}(x \notin B) \\
&\Leftrightarrow \forall B \in \mathcal{B}(x \in A \wedge x \notin B) \\
&\Leftrightarrow \forall B \in \mathcal{B}(x \in A - B) \\
&\Leftrightarrow x \in \bigcap_{B \in \mathcal{B}} (A - B),
\end{aligned}$$

which of course shows the desired result. $\square$

For intersections, we claim that

$$A - \bigcap_{B \in \mathcal{B}} B = \bigcup_{B \in \mathcal{B}} (A - B).$$

Proof. Similarly, we show this with a series of logically equivalent statements. For any x we have

$$\begin{aligned} x \in A - \bigcap_{B\in\mathcal{B}} B &\Leftrightarrow x \in A \wedge x \notin \bigcap_{B\in\mathcal{B}} B \\ &\Leftrightarrow x \in A \wedge \neg\forall B \in \mathcal{B}(x \in B) \\ &\Leftrightarrow x \in A \wedge \exists B \in \mathcal{B}(x \notin B) \\ &\Leftrightarrow \exists B \in \mathcal{B}(x \in A \wedge x \notin B) \\ &\Leftrightarrow \exists B \in \mathcal{B}(x \in A - B) \\ &\Leftrightarrow x \in \bigcup_{B\in\mathcal{B}} (A - B)\,, \end{aligned}$$

which shows the desired result. □

Exercise 1.10

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) $\{(x, y) \mid x \text{ is an integer}\}$.

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

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

(d) $\{(x, y) \mid x \text{ is not an integer and } y \text{ is an integer}\}$.

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

Solution:

(a) This is equal to the set $\mathbb{Z}\times\mathbb{R}$, which is trivial to prove.

(b) It is easy to show that this is equal to $\mathbb{R}\times(0, 1]$, where of course $(a, b]$ denotes the half-open interval $\{x \in \mathbb{R} \mid a < x \leq b\}$.

(c) We claim that this cannot be equal to the cartesian product of subsets of $\mathbb{R}$.

Proof. Let $A = \{(x, y) \mid y > x\}$ and suppose to the contrary that $A = B \times C$ where $B, C \subset \mathbb{R}$. Since $1 > 0$, we have that $(0, 1) \in A$. Then also $0 \in B$ and $1 \in C$ since $A = B \times C$. We also have that $1 \in B$ and $2 \in C$ since $2 > 1$ so that $(1, 2) \in A = B \times C$. Thus $1 \in B$ and $1 \in C$ so that $(1, 1) \in B \times C = A$, but this cannot be since it is not true that $1 > 1$. Hence we have a contradiction so that A cannot be expressed as $B \times C$. □

(d) It is trivial to show that this set is equal to $(\mathbb{R} - \mathbb{Z}) \times \mathbb{Z}$.

(e) We claim that this set cannot be expressed as the cartesian product of subsets of $\mathbb{R}$.

Proof. Let $A = \left\{(x, y) \mid x^2 + y^2 < 1\right\}$ and suppose to the contrary that $A = B \times C$ where $B, C \subset \mathbb{R}$. We then have that $(9/10)^2 + 0^2 = 81/100 + 0 = 81/100 < 1$ so that $(9/10, 0) \in A = B \times C$, and hence $9/10 \in B$ and $0 \in C$. Also $0^2 + (9/10)^2 = (9/10)^2 + 0^2 = 81/100 < 1$ so that $(0, 9/10) \in A = B \times C$, and hence $0 \in B$ and $9/10 \in C$. Hence $(9/10, 9/10) \in B \times C = A$ since $9/10$ is in both B and C. However, we have $(9/10)^2 + (9/10)^2 = 81/100 + 81/100 = 162/100 \geq 1$ so that $(9/10, 9/10)$ cannot be in A, so we have a contradiction. So it must be that A cannot be equal to $B \times C$. □

§4 The Integers and the Real Numbers

Exercise 4.1

Prove the following "laws of algebra" for $\mathbb{R}$, using only axioms (1)-(5):

(a) If $x + y = x$, then $y = 0$.
(b) $0 \cdot x = 0$. [Hint: Compute $(x + 0) \cdot x$.]
(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:

Lemma 4.1.1. *$x + y = x + z$ if and only if $y = z$.*

Proof. ($\Leftarrow$) Clearly if $y = z$ then $x + y = x + z$ since the $+$ operation is a function.
($\Rightarrow$) If $x + y = x + z$ then we have

$$
\begin{aligned}
y &= y + 0 && \text{(by (3))} \\
&= 0 + y && \text{(by (2))} \\
&= (x + (-x)) + y && \text{(by (4))} \\
&= (-x + x) + y && \text{(by (2))} \\
&= -x + (x + y) && \text{(by (1))} \\
&= -x + (x + z) && \text{(by what was just shown for } (\Leftarrow)) \\
&= (-x + x) + z && \text{(by (1))} \\
&= (x + (-x)) + z && \text{(by (2))} \\
&= 0 + z && \text{(by (4))} \\
&= z + 0 && \text{(by (2))} \\
&= z && \text{(by (3))}
\end{aligned}
$$

as desired. □

Lemma 4.1.2. *If $x \neq 0$ then $x \cdot y = x \cdot z$ if and only if $y = z$.*

Proof. ($\Leftarrow$) Clearly if $y = z$ then $x \cdot y = x \cdot z$ since the $\cdot$ operation is a function.
($\Rightarrow$) If $x \cdot y = x \cdot z$ then we have

$$
\begin{aligned}
y &= y \cdot 1 && \text{(by (3))} \\
&= 1 \cdot y && \text{(by (2))}
\end{aligned}
$$

$$
\begin{aligned}
&= \left(x \cdot \frac{1}{x}\right) \cdot y && \text{(by (4), noting that } x \neq 0\text{)} \\
&= \left(\frac{1}{x} \cdot x\right) \cdot y && \text{(by (2))} \\
&= \frac{1}{x} \cdot (x \cdot y) && \text{(by (1))} \\
&= \frac{1}{x} \cdot (x \cdot z) && \text{(by what was just shown for } (\Leftarrow)\text{)} \\
&= \left(\frac{1}{x} \cdot x\right) \cdot z && \text{(by (1))} \\
&= \left(x \cdot \frac{1}{x}\right) \cdot z && \text{(by (2))} \\
&= 1 \cdot z && \text{(by (4))} \\
&= z \cdot 1 && \text{(by (2))} \\
&= z && \text{(by (3))}
\end{aligned}
$$

as desired. □

Lemma 4.1.3. $1/(yz) = 1/(zy)$ *if* $y, z \neq 0$.

Proof. We have $(zy) \cdot 1/(yz) = (yz) \cdot 1/(yz) = 1$ by (2) followed by (4) so that $1/(yz)$ is a reciprocal of zy. Since this reciprocal is unique, however, it must be that $1/(yz) = 1/(zy)$ as desired. □

Main Problem.

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

Proof. Clearly by (3) we have $x + 0 = x = x + y$ so that it has to be that $y = 0$ by Lemma 4.1.1. □

(b) $0 \cdot x = 0$. [Hint: Compute $(x + 0) \cdot x$.]

Proof. We have

$$
\begin{aligned}
x \cdot x + 0 \cdot x &= x \cdot x + x \cdot 0 && \text{(since } 0 \cdot x = x \cdot 0 \text{ by (2))} \\
&= x \cdot (x + 0) && \text{(by (5))} \\
&= x \cdot x\,. && \text{(since } x + 0 = x \text{ by (3))}
\end{aligned}
$$

Thus it must be that $0 \cdot x = 0$ by part (a). □

(c) $-0 = 0$.

Proof. By (4) we have $0 + (-0) = 0$ so that it has to be that $-0 = 0$ by part (a). □

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

Proof. We have

$$
\begin{aligned}
-(-x) &= -(-x) + 0 && \text{(by (3))} \\
&= -(-x) + (x + (-x)) && \text{(by (4))} \\
&= -(-x) + ((-x) + x) && \text{(by (2))}
\end{aligned}
$$

$$
\begin{aligned}
&= (-(-x) + (-x)) + x && \text{(by (1))} \\
&= ((-x) + (-(-x))) + x && \text{(by (2))} \\
&= 0 + x && \text{(by (4))} \\
&= x + 0 && \text{(by (2))} \\
&= x && \text{(by (3))}
\end{aligned}
$$

as desired. □

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

Proof. First we have

$$
\begin{aligned}
x(-y) &= x(-y) + 0 && \text{(by (3))} \\
&= x(-y) + (xy + (-(xy)) && \text{(by (4))} \\
&= (x(-y) + xy) + (-(xy)) && \text{(by (1))} \\
&= x(-y + y) + (-(xy)) && \text{(by (5))} \\
&= x(y + (-y)) + (-(xy)) && \text{(by (2))} \\
&= x \cdot 0 + (-(xy)) && \text{(by (4))} \\
&= 0 \cdot x + (-(xy)) && \text{(by (2))} \\
&= 0 + (-(xy)) && \text{(by part(b))} \\
&= -(xy) + 0 && \text{(by (2))} \\
&= -(xy)\,. && \text{(by (3))}
\end{aligned}
$$

We also have

$$
\begin{aligned}
(-x)y &= y(-x) && \text{(by (2))} \\
&= -(yx) && \text{(by what was just shown)} \\
&= -(xy) && \text{(by (2))}
\end{aligned}
$$

so that the result follows since equality is transitive. □

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

Proof. We have

$$
\begin{aligned}
(-1)x &= -(1 \cdot x) && \text{(by part(e))} \\
&= -(x \cdot 1) && \text{(by (2))} \\
&= -x && \text{(since } x \cdot 1 = x \text{ by (3))}
\end{aligned}
$$

as desired. □

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

Proof. We have

$$
\begin{aligned}
x(y - z) &= x(y + (-z)) && \text{(by the definition of subtraction)} \\
&= xy + x(-z) && \text{(by (5))} \\
&= xy + (-(xz)) && \text{(by part(e))} \\
&= xy - xz && \text{(by the definition of subtraction)}
\end{aligned}
$$

as desired. □

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

Proof. We have

$$\begin{aligned}
-(x+y) &= (-1)(x+y) && \text{(by part (f))}\\
&= (-1)x + (-1)y && \text{(by (5))}\\
&= -x + (-y) && \text{(by part (f) twice)}\\
&= -x - y && \text{(by the definition of subtraction)}
\end{aligned}$$

and

$$\begin{aligned}
-(x-y) &= -(x+(-y)) && \text{(by the definition of subtraction)}\\
&= -x - (-y)) && \text{(by what was just shown)}\\
&= -x + (-(-y)) && \text{(by the definition of subtraction)}\\
&= -x + y && \text{(by part (d))}
\end{aligned}$$

as desired. □

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

Proof. By (3) we have $x \cdot 1 = x = x \cdot y$ so that it has to be that $y = 1$ by Lemma 4.1.2, noting that this applies since $x \neq 0$. □

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

Proof. By the definition of division we have $x/x = x \cdot (1/x) = 1$ by (4) since $x \neq 0$ and $1/x$ is defined as the reciprocal (i.e. the multiplicative inverse) of x. □

(k) $x/1 = x$.

Proof. First, we have by (4) that $1 \cdot (1/1) = 1$, where $1/1$ is the reciprocal of 1. We also have that $1 \cdot (1/1) = (1/1) \cdot 1 = 1/1$ by (2) and (3). Therefore $1/1 = 1 \cdot (1/1) = 1$ so that 1 is its own reciprocal. Then, by the definition of division, we have $x/1 = x \cdot (1/1) = x \cdot 1 = x$ by (3). □

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

Proof. Suppose that $x \neq 0$ and $y \neq 0$. Also suppose to the contrary that $xy = 0$. Since $y \neq 0$ it follows from (4) that $1/y$ exists. So, we have $(xy) \cdot (1/y) = 0 \cdot (1/y) = 0$ by part (b). We also have

$$\begin{aligned}
(xy) \cdot \frac{1}{y} &= x\left(y \cdot \frac{1}{y}\right) && \text{(by (1))}\\
&= x \cdot 1 && \text{(by (4))}\\
&= x && \text{(by (3))}
\end{aligned}$$

so that $x = (xy) \cdot (1/y) = 0$, which is a contradiction since we supposed that $x \neq 0$. Hence it must be that $xy \neq 0$ as desired. □

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

Proof. We have

$$
\begin{aligned}
(yz)\left(\frac{1}{y}\cdot\frac{1}{z}\right) &= (yz)\left(\frac{1}{z}\cdot\frac{1}{y}\right) && \text{(by (2))}\\
&= \left((yz)\cdot\frac{1}{z}\right)\frac{1}{y} && \text{(by (1))}\\
&= \left(y\left(z\cdot\frac{1}{z}\right)\right)\frac{1}{y} && \text{(by (1))}\\
&= (y\cdot 1)\frac{1}{y} && \text{(by (4))}\\
&= y\cdot\frac{1}{y} && \text{(by (3))}\\
&= 1 && \text{(by (4))}
\end{aligned}
$$

so that $(1/y)(1/z)$ is a multiplicative inverse of yz. Since this inverse is *unique* by (4), however, it has to be that $(1/y)(1/z) = 1/(yz)$ as desired. □

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

Proof. We have

$$
\begin{aligned}
\frac{x}{y}\cdot\frac{w}{z} &= \left(x\cdot\frac{1}{y}\right)\left(w\cdot\frac{1}{z}\right) && \text{(by the definition of division)}\\
&= \left(x\cdot\frac{1}{y}\right)\left(\frac{1}{z}\cdot w\right) && \text{(by (2))}\\
&= \left(\left(x\cdot\frac{1}{y}\right)\frac{1}{z}\right)w && \text{(by (1))}\\
&= \left(x\left(\frac{1}{y}\cdot\frac{1}{z}\right)\right)w && \text{(by (1))}\\
&= \left(x\cdot\frac{1}{yz}\right)w && \text{(by part (m) since } y, z \neq 0\text{)}\\
&= \left(\frac{1}{yz}\cdot x\right)w && \text{(by (2))}\\
&= \frac{1}{yz}(xw) && \text{(by (1))}\\
&= (xw)\frac{1}{yz} && \text{(by (2))}\\
&= \frac{xw}{yz} && \text{(by the definition of division)}
\end{aligned}
$$

as desired. □

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

Proof. We have

$$
\begin{aligned}
\frac{x}{y} + \frac{w}{z} &= \frac{x}{y}\cdot 1 + \frac{w}{z}\cdot 1 && \text{(by (3))}\\
&= \frac{x}{y}\cdot\frac{z}{z} + \frac{w}{z}\cdot\frac{y}{y} && \text{(by part (j))}
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{xz}{yz} + \frac{wy}{zy} && \text{(by part(n))} \\
&= (xz)\frac{1}{yz} + (wy)\frac{1}{zy} && \text{(by the definition of division)} \\
&= (xz)\frac{1}{yz} + (wy)\frac{1}{yz} && \text{(by Lemma 4.1.3)} \\
&= \frac{1}{yz}(xz) + \frac{1}{yz}(wy) && \text{(by (2))} \\
&= \frac{1}{yz}\left(xz + wy\right) && \text{(by (5))} \\
&= \left(xz + wy\right)\frac{1}{yz} && \text{(by (2))} \\
&= \frac{xz + wy}{yz} && \text{(by the definition of division)}
\end{aligned}
$$

as desired. □

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

Proof. Suppose that $x \neq 0$ but $1/x = 0$. Then we first have that $x \cdot (1/x) = x \cdot 0 = 0 \cdot x = 0$ by (2) and part (b). However, we also have $x \cdot (1/x) = 1$ by (4). Hence we have $0 = x \cdot (1/x) = 1$, which is a contradiction since we know that 0 and 1 are distinct by (3). So, if we accept that $x \neq 0$, then it must be that $1/x \neq 0$ also. □

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

Proof. We have

$$
\begin{aligned}
\frac{w}{z} \cdot \frac{z}{w} &= \frac{wz}{zw} && \text{(by part (n) since } w, z \neq 0) \\
&= (wz)\frac{1}{zw} && \text{(by the definition of division)} \\
&= (wz)\frac{1}{wz} && \text{(by Lemma 4.1.3 since } w, z \neq 0) \\
&= 1 && \text{(by (4))}
\end{aligned}
$$

so that by definition z/w is the reciprocal of w/z. Since this is unique by (4) we then have $z/w = 1/(w/z)$ as desired. □

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

Proof. We have

$$
\begin{aligned}
\frac{x/y}{w/z} &= \frac{x}{y} \cdot \frac{1}{w/z} && \text{(by the definition of division)} \\
&= \frac{x}{y} \cdot \frac{z}{w} && \text{(by part (q) since } w, z \neq 0) \\
&= \frac{xz}{yw} && \text{(by part (n) since } y, w \neq 0)
\end{aligned}
$$

as desired. □

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

Proof. We have

$$
\begin{aligned}
\frac{ax}{y} &= (ax) \cdot \frac{1}{y} && \text{(by the definition of division)} \\
&= a\left(x \cdot \frac{1}{y}\right) && \text{(by (1))} \\
&= a \cdot \frac{x}{y} && \text{(by the definition of division)}
\end{aligned}
$$

as desired. □

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

Proof. We have

$$
\begin{aligned}
\frac{-x}{y} &= (-x) \cdot \frac{1}{y} && \text{(by the definition of division)} \\
&= ((-1)x) \cdot \frac{1}{y} && \text{(by part (f))} \\
&= (-1)\left(x \cdot \frac{1}{y}\right) && \text{(by (1))} \\
&= (-1)\frac{x}{y} && \text{(by the definition of division)} \\
&= -\left(\frac{x}{y}\right). && \text{(by part (f))}
\end{aligned}
$$

Now, we have $(-1)(-1) = -(-1) = 1$ by parts (f) and (d) so that -1 is its own reciprocal, since the reciprocal is unique, i.e. $1/(-1) = -1$. We also have

$$
\begin{aligned}
\frac{-x}{y} &= (-x) \cdot \frac{1}{y} && \text{(by the definition of division)} \\
&= ((-1)x) \cdot \frac{1}{y} && \text{(by part (f))} \\
&= (x(-1)) \cdot \frac{1}{y} && \text{(by (2))} \\
&= x\left((-1)\frac{1}{y}\right) && \text{(by (1))} \\
&= x\left(\frac{1}{-1} \cdot \frac{1}{y}\right) && \text{(by what was just shown above)} \\
&= x\frac{1}{(-1)y} && \text{(part (m) since } y \neq 0) \\
&= x\frac{1}{-y} && \text{(by part (f))}
\end{aligned}
$$

so that $-(x/y) = (-x)/y = x/(-y)$ as desired. □

Exercise 4.2

Prove the following "laws of inequalities" for $\mathbb{R}$, using axioms (1)-(6) along with the results 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:

Lemma 4.2.1. $x + x = 2x$ *for any real* x.

Proof. We simply have

$$\begin{aligned} x + x &= x \cdot 1 + x \cdot 1 && \text{(by (3))} \\ &= x(1 + 1) && \text{(by (5))} \\ &= x \cdot 2 && \text{(since 2 is defined as } 1 + 1) \\ &= 2x && \text{(by (2))} \end{aligned}$$

as desired. □

Main Problem.

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

Proof. We have

$$\begin{aligned} x + w &> y + w && \text{(by (6) since } x > y) \\ &= w + y && \text{(by (2))} \\ &> z + y && \text{(by (6) since } w > z) \\ &= y + z && \text{(by (2))} \end{aligned}$$

as desired. □

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

Proof. First we have

$$\begin{aligned} x + y &> 0 + y && \text{(by (6) since } x > 0) \\ &= y + 0 && \text{(by (2))} \\ &= y && \text{(by (3))} \\ &> 0\,. \end{aligned}$$

Also

$$\begin{aligned} x \cdot y &> 0 \cdot y && \text{(by (6) since } x > 0 \text{ and } y > 0) \\ &= 0 && \text{(by Exercise 4.1b)} \end{aligned}$$

as desired. □

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

Proof. (⇒) Suppose that $x > 0$. Then we have

$$\begin{aligned}
-x &= -x + 0 && \text{(by (3))}\\
&= 0 + (-x) && \text{(by (2))}\\
&< x + (-x) && \text{(by (6) since } 0 < x)\\
&= 0\,. && \text{(by (4))}
\end{aligned}$$

(⇐) Suppose now that $-x < 0$. Then we have

$$\begin{aligned}
x &= x + 0 && \text{(by (3))}\\
&= 0 + x && \text{(by (2))}\\
&> -x + x && \text{(by (6) since } 0 > -x)\\
&= x + (-x) && \text{(by (2))}\\
&= 0 && \text{(by (4))}
\end{aligned}$$

as desired. □

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

Proof. (⇒) Suppose that $x > y$. Then we have

$$\begin{aligned}
-y &= -y + 0 && \text{(by (3))}\\
&= -y + (x + (-x)) && \text{(by (4))}\\
&= (x + (-x)) + (-y) && \text{(by (2))}\\
&= x + (-x + (-y)) && \text{(by (1))}\\
&> y + (-x + (-y)) && \text{(by (6) since } x > y)\\
&= y + (-y + (-x)) && \text{(by (2))}\\
&= (y + (-y)) + (-x) && \text{(by (1))}\\
&= 0 + (-x) && \text{(by (4))}\\
&= -x + 0 && \text{(by (2))}\\
&= -x\,. && \text{(by (3))}
\end{aligned}$$

(⇐) Now suppose that $-x < -y$. Then we have

$$\begin{aligned}
x &= x + 0 && \text{(by (3))}\\
&= x + (y + (-y)) && \text{(by (4))}\\
&= (y + (-y)) + x && \text{(by (2))}\\
&= (-y + y) + x && \text{(by (2))}\\
&= -y + (y + x) && \text{(by (1))}\\
&> -x + (y + x) && \text{(by (6) since } -y > -x)\\
&= -x + (x + y) && \text{(by (2))}\\
&= (-x + x) + y && \text{(by (1))}\\
&= (x + (-x)) + y && \text{(by (2))}\\
&= 0 + y && \text{(by (4))}\\
&= y + 0 && \text{(by (2))}\\
&= y && \text{(by (3))}
\end{aligned}$$

as desired. □

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

Proof. First, by Exercise 4.1d, we have $-(-z) = z < 0$ so that $-z > 0$ by part (c). Then, since $x > y$, it follows from (6) that

$$\begin{aligned} x(-z) &> y(-z) \\ -(xz) &> -(yz) && \text{(by Exercise 4.1e applied to both sides)} \\ xz &< yz && \text{(by part (d))} \end{aligned}$$

as desired. ☐

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

Proof. Since $x \neq 0$ we either have that $x > 0$ or $x < 0$ since the $<$ relation is an order (in particular a linear order since this is part of the definition of order in this text). If $x > 0$ then we have $x^2 = x \cdot x > 0 \cdot x = 0$ by (6) (since $x > 0$) and Exercise 4.1b. If $x < 0$ then we have $0 = 0 \cdot x < x \cdot x = x^2$ by part (e) (since $0 > x$) and Exercise 4.1b. Together these show the desired result. ☐

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

Proof. By (4) we know that $1 \neq 0$ so that $1^2 > 0$ by part (f). However, we have $1^2 = 1 \cdot 1 = 1$ by (3). Hence $1 = 1^2 > 0$. It then follows from part (c) that $-1 < 0$ so that we have $-1 < 0 < 1$ as desired. ☐

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

Proof. ($\Rightarrow$) Suppose that $xy > 0$. It cannot be that $x = 0$, for then we would have $0 = 0 \cdot y = xy > 0$ by Exercise 4.1b, which is impossible by the definition of an order. Hence we have $x \neq 0$, and an analogous argument shows that $y \neq 0$ as well. We then have the following:

Case: $x > 0$. Suppose that $y < 0$. Then, by part (e) and Exercise 4.1b, we have $xy < 0 \cdot y = 0$ since $x > 0$ and $y < 0$, which contradicts our initial supposition. Thus, since we know that $y \neq 0$, it has to be that $y > 0$ as well.

Case: $x < 0$. Suppose that $y > 0$. Then, by (6) and Exercise 4.1b, we have $0 = 0 \cdot y > xy$ since $0 > x$ and $y > 0$, which again contradicts the initial supposition. So it must be that $y < 0$ also since $y \neq 0$.

Therefore in every case either both x and y are positive or they are both negative. Since $x \neq 0$, these cases are exhaustive so that this shows the result.

($\Leftarrow$) Suppose that either $x > 0, y > 0$ or $x < 0, y < 0$. In the case where both $x > 0$ and $y > 0$ we clearly have $xy > 0 \cdot y = 0$ by (6) and Exercise 4.1b. In the other case in which $x < 0$ and $y < 0$ we have $0 = 0 \cdot y < xy$ by part (e) and Exercise 4.1b since $0 > x$ and $y < 0$. Hence $xy > 0$ in both cases. ☐

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

Proof. First, it cannot be that $1/x = 0$ because then we would have $1 = x(1/x) = x \cdot 0 = 0 \cdot x = 0$ by (4), (2), and Exercise 4.1b. This is clearly a contradiction since we know that $1 \neq 0$ by (3). Hence $1/x \neq 0$. Now suppose that $1/x < 0$ so that $1 = x(1/x) < 0 \cdot (1/x) = 0$ by part (e) since $x > 0$ and $1/x < 0$, and we have also used Exercise 4.1b. This is also a contradiction since it was proved in part (g) that $1 > 0$. Hence the only remaining possibility is that $1/x > 0$ as desired. ☐

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

Proof. First, since the order is transitive, we have $x, y > 0$. It then follows from part (i) that $1/x, 1/y > 0$. Then $(1/x)(1/y) > 0$ by part (h). We then have

$$\begin{aligned}
\frac{1}{x} &= \frac{1}{x} \cdot 1 && \text{(by (3))} \\
&= \frac{1}{x}\left(y \cdot \frac{1}{y}\right) && \text{(by (4))} \\
&= \left(\frac{1}{x} \cdot y\right)\frac{1}{y} && \text{(by (1))} \\
&= \left(y \cdot \frac{1}{x}\right)\frac{1}{y} && \text{(by (2))} \\
&= y\left(\frac{1}{x} \cdot \frac{1}{y}\right) && \text{(by (1))} \\
&< x\left(\frac{1}{x} \cdot \frac{1}{y}\right) && \text{(by (6) since } y < x \text{ and } (1/x)(1/y) > 0) \\
&= \left(x \cdot \frac{1}{x}\right)\frac{1}{y} && \text{(by (1))} \\
&= 1 \cdot \frac{1}{y} && \text{(by (4))} \\
&= \frac{1}{y} \cdot 1 && \text{(by (2))} \\
&= \frac{1}{y} && \text{(by (3))}
\end{aligned}$$

as desired. □

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

Proof. First, we know by part (g) that $1 > 0$ so that

$$\begin{aligned}
2 &= 1 + 1 && \text{(by the definition of 2)} \\
&> 0 + 1 && \text{(by (6) since } 1 > 0\text{)} \\
&= 1 + 0 && \text{(by (2))} \\
&= 1 && \text{(by (3))} \\
&> 0\,. && \text{(by part (g))}
\end{aligned}$$

To summarize, $0 < 1 < 2$. It then follows from part (i) that $1/2 > 0$. We then have

$$\begin{aligned}
x &< y \\
x + x &< x + y && \text{(by (6))} \\
2x &< x + y && \text{(by Lemma 4.2.1)} \\
(2x)\frac{1}{2} &< (x+y)\frac{1}{2} && \text{(by (6) since } 1/2 > 0) \\
(x \cdot 2)\frac{1}{2} &< \frac{x+y}{2} && \text{(by (2) and the definition of division)} \\
x\left(2 \cdot \frac{1}{2}\right) &< \frac{x+y}{2} && \text{(by (1))} \\
x \cdot 1 &< \frac{x+y}{2} && \text{(by (4))}
\end{aligned}$$

$$x < \frac{x+y}{2}\,. \qquad \text{(by (3))}$$

Similarly, we have

$$\begin{aligned}
x &< y \\
x + y &< y + y && \text{(by (6))} \\
x + y &< 2y && \text{(by Lemma 4.2.1)} \\
(x+y)\frac{1}{2} &< (2y)\frac{1}{2} && \text{(by (6) since } 1/2 > 0\text{)} \\
\frac{x+y}{2} &< (y \cdot 2)\frac{1}{2} && \text{(by the definition of division and (2))} \\
\frac{x+y}{2} &< y\left(2 \cdot \frac{1}{2}\right) && \text{(by (1))} \\
\frac{x+y}{2} &< y \cdot 1 && \text{(by (4))} \\
\frac{x+y}{2} &< y\,. && \text{(by (3))}
\end{aligned}$$

This shows that $x < (x+y)/2 < y$ as desired. □

Exercise 4.3

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

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

Solution:

(a) We must show that $\bigcap_{A \in \mathcal{A}} A$ is inductive.

Proof. First, consider any $A \in \mathcal{A}$. Then, since A is inductive, $1 \in A$. Since A was arbitrary, this shows that $1 \in \bigcap_{A \in \mathcal{A}} A$. Now suppose that $x \in \bigcap_{A \in \mathcal{A}} A$ and again consider arbitrary $A \in \mathcal{A}$. Then $x \in A$ so that $x + 1 \in A$ also since A is inductive. Since A was arbitrary, this shows that $x + 1 \in \bigcap_{A \in \mathcal{A}} A$. Hence by definition $\bigcap_{A \in \mathcal{A}} A$ is inductive. □

(b)

Proof. Let $\mathcal{A}$ be the collection of all inductive sets of $\mathbb{R}$ so that by definition $\mathbb{Z}_+ = \bigcap_{A \in \mathcal{A}} A$. It then follows immediately from part (a) that $\mathbb{Z}_+$ is inductive since $\mathcal{A}$ is a collection of inductive sets. This shows property (1).

Now suppose that A is an inductive set of positive integers. That is, A is inductive and $A \subset \mathbb{Z}_+$. Consider any $x \in \mathbb{Z}_+ = \bigcap_{B \in \mathcal{A}} B$, where again $\mathcal{A}$ is the the collection of all inductive subsets of $\mathbb{R}$. Clearly we have that $A \subset \mathbb{Z}_+ \subset \mathbb{R}$ so that $A \in \mathcal{A}$ since A is an inductive subset of $\mathbb{R}$. Hence $x \in A$ (since $x \in \bigcap_{B \in \mathcal{A}} B$ and $A \in \mathcal{A}$) so that $\mathbb{Z}_+ \subset A$ since x was arbitrary. This shows that $A = \mathbb{Z}_+$ as desired since also $A \subset \mathbb{Z}_+$. This shows property (2). □

Exercise 4.4

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

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

Solution:

(a)

Proof. Let A be the set of integers such that the hypothesis is true. Clearly the result is then shown if we can prove that $A = \mathbb{Z}_+$. So first, clearly $1 \in A$ since the set $\{1\}$ has only a single nonempty subset, i.e. $\{1\}$ itself, in which 1 is clearly the largest element. Now suppose that $n \in A$ so that every nonempty subset of $S_{n+1} = \{1, \ldots, n\}$ has a largest element. Consider any nonempty subset B of $S_{n+2} = \{1, \ldots, n+1\}$, noting that $S_{n+2} = S_{n+1} \cup \{n+1\}$.

Case: $n+1 \in B$. Then, for any other $k \in B$, $k \in S_{n+2}$ so that either $k = n+1$ or $k \in S_{n+1}$ so that $k < n+1$ by the definition of S_{n+1}. Thus in either case $k \leq n+1$ so that $n+1$ is the largest element of B since k was arbitrary.

Case: $n+1 \notin B$. Then clearly $B \subset S_{n+1}$ so that B has a largest element by the induction hypothesis since B is nonempty.

Hence in either case B has a largest element so that $n+1 \in A$ since B was an arbitrary nonempty subset of $S_{n+2} = \{1, \ldots, n+1\}$. This shows that A is an inductive set of positive integers so that $A = \mathbb{Z}_+$ as desired by the Principle of Induction. $\square$

(b) There could be nonempty subsets of $\mathbb{Z}_+$ that are *not* subsets of $S_{n+1} = \{1, \ldots, n\}$ for any $n \in \mathbb{Z}_+$, in which cases the hypothesis of part (a) is not satisfied so that the conclusion does not necessarily apply. In fact, $\mathbb{Z}_+$ itself is an example of such a set where both the hypothesis and the conclusion are false.

Exercise 4.5

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

(a) $a, b \in \mathbb{Z}_+ \Rightarrow a + b \in \mathbb{Z}_+$. [Hint: Show that given $a \in \mathbb{Z}_+$, the set $X = \{x \mid x \in \mathbb{R} \text{ and } a + x \in \mathbb{Z}_+\}$ is inductive.]

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

(c) Show that $a \in \mathbb{Z}_+ \Rightarrow a - 1 \in \mathbb{Z}_+ \cup \{0\}$. [Hint: Let $X = \{x \mid x \in \mathbb{R} \text{ and } x - 1 \in \mathbb{Z}_+ \cup \{0\}\}$; show that X is inductive.]

(d) $c, d \in \mathbb{Z} \Rightarrow c + d \in \mathbb{Z}$ and $c - d \in \mathbb{Z}$. [Hint: Prove it first for $d = 1$.]

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

Solution:

Lemma 4.5.1. *If $x \in \mathbb{Z}$ then $-x \in \mathbb{Z}$.*

Proof. Let $\mathbb{Z}_- = \{-x \mid x \in \mathbb{Z}_+\}$ so that by definition $\mathbb{Z} = \mathbb{Z}_+ \cup \{0\} \cup \mathbb{Z}_-$. Suppose that $x \in \mathbb{Z}$ so that $x \in \mathbb{Z}_+ \cup \{0\} \cup \mathbb{Z}_-$.

Case: $x \in \mathbb{Z}_+$. Then $-x \in \mathbb{Z}_-$ by definition.

Case: $x = 0$. Then by Exercise 4.1c we have $-x = -0 = 0 \in \{0\}$.

Case: $x \in \mathbb{Z}_-$. Then by definition there is a $y \in \mathbb{Z}_+$ such that $x = -y$. Then $-x = -(-y) = y \in \mathbb{Z}_+$ by Exercise 4.1d.

Hence in all cases either $-x \in \mathbb{Z}_+$, $-x \in \{0\}$, or $-x \in \mathbb{Z}_-$ so that $-x \in \mathbb{Z}_+ \cup \{0\} \cup \mathbb{Z}_- = \mathbb{Z}$ as desired. $\square$

Main Problem.

(a)

Proof. Consider any $a \in \mathbb{Z}_+$ and define $X_a = \{x \in \mathbb{R} \mid a + x \in \mathbb{Z}_+\}$. We show that X_a is inductive. First, since $a \in \mathbb{Z}_+$ we have that $a+1 \in \mathbb{Z}_+$ since $\mathbb{Z}_+$ is inductive. Hence $1 \in X_a$ by definition. Now suppose that $x \in X_a$ so that $a + x \in \mathbb{Z}_+$. Then we have $a + (x+1) = (a+x)+1 \in \mathbb{Z}_+$ since $a + x \in \mathbb{Z}_+$ and $\mathbb{Z}_+$ is inductive. This shows by definition that $x + 1 \in X_a$ and therefore that X_a is inductive. It follows that $\mathbb{Z}_+ \subset X_a$ since $\mathbb{Z}_+$ is defined as the intersection of all inductive subsets of reals, of which X_a is one.

Therefore, for any $a, b \in \mathbb{Z}_+$, we have that $b \in X_a$ since $\mathbb{Z}_+ \subset X_a$. Thus by definition $a + b \in \mathbb{Z}_+$ as desired $\square$

(b)

Proof. Consider any $a \in \mathbb{Z}_+$ and define $X_a = \{x \in \mathbb{R} \mid a \cdot x \in \mathbb{Z}_+\}$. We show that X_a is inductive. To this end, we first have that $a \cdot 1 = a \in \mathbb{Z}_+$ so that $1 \in X_a$ by definition. Now suppose that $x \in X_a$ so that $ax \in \mathbb{Z}_+$. Then we have $a \cdot (x+1) = a \cdot x + a \cdot 1 = ax + a \in \mathbb{Z}_+$ by part (a) since we know both ax and a are in $\mathbb{Z}_+$. Hence $x + 1 \in X_a$ by definition. This shows that X_a is inductive so that again $\mathbb{Z}_+ \subset X_a$.

Hence for any $a, b \in \mathbb{Z}_+$ we have that $b \in X_a$ since $\mathbb{Z}_+ \subset X_a$. It then follows by definition that $a \cdot b \in \mathbb{Z}_+$ as desired. $\square$

(c)

Proof. Let $X = \{x \in \mathbb{R} \mid x - 1 \in \mathbb{Z}_+ \cup \{0\}\}$, which we show is inductive. First, we have $1 - 1 = 1 + (-1) = 0$ so that clearly $1 \in \mathbb{Z}_+ \cup \{0\}$ and hence $1 \in X$. Now suppose that $x \in X$ so that $x - 1 \in \mathbb{Z}_+ \cup \{0\}$.

Case: $x - 1 \in \{0\}$. Then it must be that $x - 1 = 0$, which clearly implies $x = 1 \in \mathbb{Z}_+$ since $\mathbb{Z}_+$ is inductive. Then $(x+1) - 1 = x + (1-1) = x + 0 = x \in \mathbb{Z}_+$ so that $(x+1) - 1 \in \mathbb{Z}_+ \cup \{0\}$ and therefore $x + 1 \in X$.

Case: $x - 1 \in \mathbb{Z}_+$. Then $(x+1) - 1 = x + (1-1) = x + ((-1)+1) = (x-1)+1 \in \mathbb{Z}_+$ since $x - 1 \in \mathbb{Z}_+$ and $\mathbb{Z}_+$ is inductive. Thus clearly $(x+1) - 1 \in \mathbb{Z}_+ \cup \{0\}$ so that $x + 1 \in X$ by definition.

Hence in both cases $x + 1 \in X$, which shows that X is inductive, and so $\mathbb{Z}_+ \subset X$. Therefore, for any $a \in \mathbb{Z}_+$, we have that also $x \in X$ since $\mathbb{Z}_+ \subset X$. Then, by the definition of X, it follows that $a - 1 \in \mathbb{Z}_+ \cup \{0\}$ as desired. $\square$

(d)

Proof. First we show that the set $X_c = \{x \in \mathbb{R} \mid c + x \in \mathbb{Z} \text{ and } c - x \in \mathbb{Z}\}$ is inductive for any $c \in \mathbb{Z}$. So consider any c and b in $\mathbb{Z}$ so that $c, b \in \mathbb{Z}_+ \cup \{0\} \cup \mathbb{Z}_-$.

Case: $b \in \mathbb{Z}_+$. Then $b + 1 \in \mathbb{Z}_+$ since $\mathbb{Z}_+$ is inductive and $b - 1 \in \mathbb{Z}_+ \cup \{0\}$ by part (c).

Case: $b = 0$. Then $b + 1 = 0 + 1 = 1 \in \mathbb{Z}_+$ since it is inductive, and $b - 1 = 0 - 1 = -1 \in \mathbb{Z}_-$ since $1 \in \mathbb{Z}_+$.

Case: $b \in \mathbb{Z}_-$. Then $b = -a$ for $a \in \mathbb{Z}_+$, and we then have that $a + 1 \in \mathbb{Z}_+$ since $\mathbb{Z}_+$ is inductive. Hence $b - 1 = -a - 1 = -(a+1) \in \mathbb{Z}_-$. We also have that $a - 1 \in \mathbb{Z}_+ \cup \{0\}$ by part (c), from which it is trivial to show that $-(a-1) \in \mathbb{Z}_- \cup \{0\}$. Therefore $b + 1 = -a + 1 = -(a-1) \in \mathbb{Z}_- \cup \{0\}$.

Thus in all cases we have that $b+1$ and $b-1$ are in $\mathbb{Z}_+$ or $\{0\}$ or $\mathbb{Z}_-$ so that they are both in $\mathbb{Z}$, and so $1 \in X_b$. Note that this is the case for any $b \in \mathbb{Z}$ so that it is clearly true for c, i.e. $1 \in X_c$. Now suppose that $x \in X_c$ so that $c+x$ and $c-x$ are both in $\mathbb{Z}$. It then follows that $1 \in X_{c+x}$ and $1 \in X_{c-x}$ so that $c+(x+1) = (c+x)+1 \in \mathbb{Z}$ and $c-(x+1) = (c-x)-1 \in \mathbb{Z}$. This then shows that $x+1 \in X_c$. Hence X_c is inductive for any $c \in \mathbb{Z}$ so that $\mathbb{Z}_+ \subset X_c$.

Now consider $c, d \in \mathbb{Z}$.

Case: $d \in \mathbb{Z}_+$. Then clearly $d \in X_c$ since $\mathbb{Z}_+ \subset X_c$. Hence by definition $c+d$ and $c-d$ are both in $\mathbb{Z}$.

Case: $d = 0$. Then $c+d = c+0 = c \in \mathbb{Z}$ and $c-d = c-0 = c \in \mathbb{Z}$.

Case: $d \in \mathbb{Z}_-$. Then by definition $d = -a$ for $a \in \mathbb{Z}_+$ so that $a \in X_c$ since $\mathbb{Z}_+ \subset X_c$. Then $c+a$ and $c-a$ are both in $\mathbb{Z}$ by the definition of X_c. Hence $c+d = c+(-a) = c-a \in \mathbb{Z}$ and $c-d = c-(-a) = c+a \in \mathbb{Z}$.

Therefore we have shown that $c+d$ and $c-d$ are both integers in all cases, which is the desired result. $\square$

(e)

Proof. For any $c \in \mathbb{Z}$, define $X_c = \{x \in \mathbb{R} \mid c \cdot x \in \mathbb{Z}\}$. We first show that X_c is inductive for any such $c \in \mathbb{Z}$. We have $c \cdot 1 = c \in \mathbb{Z}$ so that $1 \in X_c$. Now suppose that $x \in X_c$ so that $c \cdot x \in \mathbb{Z}$. Then $c \cdot (x+1) = c \cdot x + c \cdot 1 = c \cdot x + c \in \mathbb{Z}$ by part (d) since both $c \cdot x$ and c are integers. This shows that X_c is inductive so that $\mathbb{Z}_+ \subset X_c$.

Now consider any $c, d \in \mathbb{Z}$.

Case: $d \in \mathbb{Z}_+$. Then $d \in X_c$ since $\mathbb{Z}_+ \subset X_c$. Thus $c \cdot d \in \mathbb{Z}$.

Case: $d = 0$. The $c \cdot d = c \cdot 0 = 0 \in \mathbb{Z}$.

Case: $d \in \mathbb{Z}_-$. Then there is an $a \in \mathbb{Z}_+$ such that $d = -a$. Hence $a \in X_c$ since $\mathbb{Z}_+ \subset X_c$, from which it follows that $c \cdot a \in \mathbb{Z}$. We then have $c \cdot d = c \cdot (-a) = -(c \cdot a) \in \mathbb{Z}$ as well by Lemma 4.5.1.

Thus in all cases $c \cdot d \in \mathbb{Z}$ as desired. $\square$

Exercise 4.6

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

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

for $n \in \mathbb{Z}_+$. (See §7 for a discussion of the process of inductive definition.) Show that for $n, m \in \mathbb{Z}_+$ and $a, b \in \mathbb{R}$,

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

These are called the ***laws of exponents***. [Hint: For fixed n, prove the formulas by induction on m.]

Solution:

The following lemma is the familiar proof by induction, which is more straightforward than having to frame everything in terms of inductive sets. Henceforth we use this whenever induction is required.

Lemma 4.6.1. *(Proof by Induction) Suppose that $P(x)$ is a statement with parameter x. Suppose also that $P(1)$ is true and that $P(x)$ implies $P(x+1)$. Then $P(n)$ is true for all $n \in \mathbb{Z}_+$.*

Proof. Define the set $X = \{x \in \mathbb{R} \mid P(x)\}$. We show that X is inductive. Clearly since $P(1)$ is true we have $1 \in X$. Now suppose that $x \in X$ so that $P(x)$ is true. Then $P(x+1)$ is also true so that $x+1 \in X$. This shows that X is inductive so that $\mathbb{Z}_+ \subset X$. So, for any positive integer n we have that $n \in X$ since $\mathbb{Z}_+ \subset X$. Therefore $P(n)$ is true. Since n was arbitrary, this shows the desired result. $\square$

Main Problem.

In what follows, suppose that $a, b \in \mathbb{R}$.

First we show that $a^n a^m = a^{n+m}$ for all $n, m \in \mathbb{Z}_+$.

Proof. Fix $n \in \mathbb{Z}_+$. We show the result by induction on m. First, we clearly have $a^n a^1 = a^n \cdot a = a^{n+1}$ by the inductive definition. Now suppose that $a^n a^m = a^{n+m}$. Then

$$\begin{aligned}
a^n a^{m+1} &= a^n \cdot (a^m \cdot a) && \text{(by the inductive definition)} \\
&= (a^n a^m) \cdot a && \text{(since multiplication is associative)} \\
&= a^{n+m} \cdot a && \text{(by the induction hypothesis)} \\
&= a^{(n+m)+1} && \text{(by the inductive definition)} \\
&= a^{n+(m+1)}, && \text{(since addition is associative)}
\end{aligned}$$

which completes the induction step. Therefore the result holds for all $m \in \mathbb{Z}_+$ by induction. $\square$

Next we show that $(a^n)^m = a^{nm}$ for all $n, m \in \mathbb{Z}_+$.

Proof. We again fix $n \in \mathbb{Z}_+$ and use induction on m. First, we have $(a^n)^1 = a^n = a^{n \cdot 1}$ by the inductive definition. Supposing now that $(a^n)^m = a^{n \cdot m}$, we have

$$\begin{aligned}
(a^n)^{m+1} &= (a^n)^m \cdot a^n && \text{(by the inductive definition)} \\
&= a^{n \cdot m} a^n && \text{(by the induction hypothesis)} \\
&= a^{n \cdot m + n} && \text{(by what was shown above)} \\
&= a^{n \cdot m + n \cdot 1} \\
&= a^{n \cdot (m+1)}. && \text{(by the distributive property)}
\end{aligned}$$

This completes the induction so that the result holds for all $m \in \mathbb{Z}_+$. $\square$

Lastly, we show that $a^m b^m = (ab)^m$ for all $m \in \mathbb{Z}_+$.

Proof. We show this by induction on m. First, we have $a^1 b^1 = ab = (ab)^1$ by the inductive definition. Now suppose that $a^m b^m = (ab)^m$ so that

$$\begin{aligned}
a^{m+1} b^{m+1} &= (a^m \cdot a)(b^m \cdot b) && \text{(by the inductive definition)} \\
&= (a \cdot a^m)(b^m \cdot b) && \text{(since multiplication is commutative)} \\
&= ((a \cdot a^m) b^m) \cdot b && \text{(since multiplication is associative)} \\
&= (a \cdot (a^m b^m)) \cdot b && \text{(since multiplication is associative)} \\
&= (a (ab)^m) \cdot b && \text{(by the induction hypothesis)} \\
&= ((ab)^m a) \cdot b && \text{(since multiplication is commutative)}
\end{aligned}$$

$$
\begin{aligned}
&= (ab)^m (ab) && \text{(since multiplication is associative)} \\
&= (ab)^{m+1}\,. && \text{(by the inductive definition)}
\end{aligned}
$$

This completes the induction. □

Exercise 4.7

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:

Lemma 4.7.1. *For any $n \in \mathbb{Z}$, $1^n = 1$.*

Proof. We show this for $n \in \mathbb{Z}_+$ by simple induction on n. First, clearly $1^1 = 1$ by the inductive definition of exponentiation. Next, if $1^n = 1$, then we have $1^{n+1} = 1^n \cdot 1 = 1^n = 1$ by the inductive definition of exponentiation and the inductive hypothesis. This completes the induction so that the result holds for all $n \in \mathbb{Z}_+$.

Clearly if $n = 0$ then, by the definition of 0 as an exponent, $1^n = 1^0 = 1$.

Lastly, if $n \in \mathbb{Z}_-$ then there is a $k \in \mathbb{Z}_+$ where $n = -k$. Then we have

$$
\begin{aligned}
1^n &= 1^{-k} \\
&= \frac{1}{1^k} && \text{(by the definition of negative exponentiation)} \\
&= \frac{1}{1} && \text{(by what was just shown by induction since } k \in \mathbb{Z}_+\text{)} \\
&= 1\,. && \text{(since 1 is its own reciprocal)}
\end{aligned}
$$

Thus the result has been shown for all the resulting cases when $n \in \mathbb{Z}$. □

Lemma 4.7.2. *$1/a^n = (1/a)^n$ for any real $a \neq 0$ and $n \in \mathbb{Z}_+$.*

Proof. We have

$$
\begin{aligned}
\left(\frac{1}{a}\right)^n a^n &= \left(\frac{1}{a} \cdot a\right)^n && \text{(by Exercise 4.6 since } n \in \mathbb{Z}_+\text{)} \\
&= 1^n && \text{(by the definition of the reciprocal)} \\
&= 1\,. && \text{(by Lemma 4.7.1)}
\end{aligned}
$$

Thus $(1/a)^n$ must be the unique reciprocal of a^n, that is $(1/a)^n = 1/a^n$ as desired. □

Lemma 4.7.3. *$a^n a^{-n} = 1$ for any real $a \neq 0$ and $n \in \mathbb{Z}_+$.*

Proof. We have

$$
\begin{aligned}
a^n a^{-n} &= a^n \left(\frac{1}{a^n}\right) && \text{(by the definition of negative exponentiation)} \\
&= a^n \left(\frac{1}{a}\right)^n && \text{(by Lemma 4.7.2)} \\
&= \left(a \cdot \frac{1}{a}\right)^n && \text{(by Exercise 4.6 since } n \in \mathbb{Z}_+\text{)}
\end{aligned}
$$

$$
\begin{aligned}
&= 1^n && \text{(by the definition of the reciprocal)} \\
&= 1 && \text{(by Lemma 4.7.1)}
\end{aligned}
$$

as desired. □

Main Problem.

First we show that $a^n a^m = a^{n+m}$ for all real $a \neq 0$ and $n, m \in \mathbb{Z}$.

Proof. Consider any real $a \neq 0$ and $n, m \in \mathbb{Z}$. We number the following cases for reference:

1. Case: $n \in \mathbb{Z}_+$.
   (a) Case: $m \in \mathbb{Z}_+$. Then the result immediately applies by Exercise 4.6.
   (b) Case: $m = 0$. Then we have $a^n a^m = a^n a^0 = a^n \cdot 1 = a^n = a^{n+0} = a^{n+m}$.
   (c) Case: $m \in \mathbb{Z}_-$. Then $m = -k$ for some $k \in \mathbb{Z}_+$.
      i. Case: $n > k$. Then $n - k > 0$ so that $n - k \in \mathbb{Z}_+$ since $n - k \in \mathbb{Z}$ by Exercise 4.5d. We then have

$$
\begin{aligned}
a^n a^m &= a^n a^{-k} \\
&= a^{n-k+k} a^{-k} && \text{(since } n = n + 0 = n - k + k\text{)} \\
&= (a^{n-k} a^k) a^{-k} && \text{(by Exercise 4.6 since } k, n - k \in \mathbb{Z}_+\text{)} \\
&= a^{n-k}(a^k a^{-k}) && \text{(since multiplication is associative)} \\
&= a^{n-k} \cdot 1 && \text{(by Lemma 4.7.3 since } k \in \mathbb{Z}_+\text{)} \\
&= a^{n-k} \\
&= a^{n+m}.
\end{aligned}
$$

      ii. Case: $n = k$. Then clearly $n + m = n - k = k - k = 0$, so that we have $a^n a^m = a^k a^{-k} = 1 = a^0 = a^{n+m}$ by Lemma 4.7.3 and the definition of 0 as an exponent.
      iii. Case: $n < k$. Then $n - k < 0$ so that $n - k \in \mathbb{Z}_-$ since $n - k \in \mathbb{Z}$ by Exercise 4.5d. Also, clearly $-n \in \mathbb{Z}_-$ since $n \in \mathbb{Z}_+$. Then we have

$$
\begin{aligned}
a^n a^m &= a^n a^{-k} \\
&= a^n a^{-k+n-n} && \text{(since } -k = -k + 0 = -k + n - n\text{)} \\
&= a^n a^{n-k-n} && \text{(since addition is commutative)} \\
&= a^n (a^{n-k} a^{-n}) && \text{(by case 3c below since } n - k, -n \in \mathbb{Z}_-\text{)} \\
&= a^n (a^{-n} a^{n-k}) && \text{(since multiplication is commutative)} \\
&= (a^n a^{-n}) a^{n-k} && \text{(since multiplication is associative)} \\
&= 1 \cdot a^{n-k} && \text{(by Lemma 4.7.3)} \\
&= a^{n-k} \\
&= a^{n+m}.
\end{aligned}
$$

2. Case: $n = 0$.
   (a) Case: $m \in \mathbb{Z}_+$. Since $a^n a^m = a^m a^n$ and $a^{n+m} = a^{m+n}$, this the same as case 1b above.
   (b) Case: $m = 0$. Then we have $a^n a^m = a^0 a^0 = 1 \cdot 1 = 1 = a^0 = a^{0+0} = a^{n+m}$.

(c) Case: $m \in \mathbb{Z}_-$. Then there is a $k \in \mathbb{Z}_+$ such that $m = -k$, and $a^n a^m = a^0 a^{-k} = 1 \cdot (1/a^k) = 1/a^k = a^{-k} = a^m = a^{0+m} = a^{n+m}$.

3. Case: $n \in \mathbb{Z}_-$.
   (a) Case: $m \in \mathbb{Z}_+$. This is the same as case 1c above.
   (b) Case: $m = 0$. This is the same as case 2c above.
   (c) Case: $m \in \mathbb{Z}_-$. Here we have that $n = -k$ and $m = -l$ for some $k, l \in \mathbb{Z}_+$. Hence we have

$$
\begin{aligned}
a^n a^m &= a^{-k} a^{-l} \\
&= \left(\frac{1}{a^k}\right)\left(\frac{1}{a^l}\right) && \text{(by the definition of negative exponents)} \\
&= \left(\frac{1}{a}\right)^k \left(\frac{1}{a}\right)^l && \text{(by Lemma 4.7.2)} \\
&= \left(\frac{1}{a}\right)^{k+l} && \text{(by Exercise 4.6 since } k, l \in \mathbb{Z}_+\text{)} \\
&= \frac{1}{a^{k+l}} && \text{(by Lemma 4.7.2)} \\
&= a^{-(k+l)} && \text{(by the definition of negative exponents)} \\
&= a^{-k-l} \\
&= a^{n+m} .
\end{aligned}
$$

Thus in all cases we have shown the result. □

Next we show that $(a^n)^m = a^{nm}$ for all real $a \neq 0$ and $n, m \in \mathbb{Z}$.

Proof. Consider any real $a \neq 0$ and $n, m \in \mathbb{Z}$. We again number the cases for reference:

1. Case: $n \in \mathbb{Z}_+$.
   (a) Case: $m \in \mathbb{Z}_+$. Then the result immediately applies by Exercise 4.6.
   (b) Case: $m = 0$. Then we have $(a^n)^m = (a^n)^0 = 1 = a^0 = a^{n \cdot 0} = a^{nm}$ by the definition of a 0 exponent.
   (c) Case: $m \in \mathbb{Z}_-$. Then there is a $k \in \mathbb{Z}_+$ such that $m = -k$. Then we have

$$
\begin{aligned}
(a^n)^m &= (a^n)^{-k} \\
&= \frac{1}{(a^n)^k} && \text{(by the definition of negative exponents)} \\
&= \frac{1}{a^{nk}} && \text{(by Exercise 4.6 since } n, k \in \mathbb{Z}_+\text{)} \\
&= a^{-(nk)} && \text{(by the definition of negative exponents)} \\
&= a^{n(-k)} \\
&= a^{nm} .
\end{aligned}
$$

2. Case: $n = 0$. Then we have $(a^n)^m = (a^0)^m = 1^m = 1 = a^0 = a^{0 \cdot m} = a^{nm}$ by the definition of 0 as an exponent and Lemma 4.7.1.
3. Case: $n \in \mathbb{Z}_-$. Then $n = -k$ for some $k \in \mathbb{Z}_+$.

(a) Case: $m \in \mathbb{Z}_+$. Then we have

$$
\begin{aligned}
(a^n)^m &= (a^{-k})^m \\
&= \left(\frac{1}{a^k}\right)^m && \text{(by the definition of negative exponents)} \\
&= \left[\left(\frac{1}{a}\right)^k\right]^m && \text{(by Lemma 4.7.2)} \\
&= \left(\frac{1}{a}\right)^{km} && \text{(by Exercise 4.6 since } k, m \in \mathbb{Z}_+\text{)} \\
&= \frac{1}{a^{km}} && \text{(by Lemma 4.7.2)} \\
&= a^{-(km)} && \text{(by the definition of negative exponents)} \\
&= a^{(-k)m} \\
&= a^{nm}\,.
\end{aligned}
$$

(b) Case: $m = 0$. The same argument as in case 1b above applies here as it does not depend on n being positive.

(c) Case: $m \in \mathbb{Z}_-$. Then $m = -l$ for some $l \in \mathbb{Z}_+$, and we have

$$
\begin{aligned}
(a^n)^m &= (a^{-k})^{-l} \\
&= \left(\frac{1}{a^k}\right)^{-l} && \text{(by the definition of negative exponents)} \\
&= \frac{1}{(1/a^k)^l} && \text{(by the definition of negative exponents)} \\
&= \frac{1}{[(1/a)^k]^l} && \text{(by Lemma 4.7.2)} \\
&= \frac{1}{(1/a)^{kl}} && \text{(by Exercise 4.6 since } k, l \in \mathbb{Z}_+\text{)} \\
&= \left(\frac{1}{1/a}\right)^{kl} && \text{(by Lemma 4.7.2)} \\
&= a^{kl} \\
&= a^{(-k)(-l)} \\
&= a^{nm}\,.
\end{aligned}
$$

Thus in all cases we have shown the result. □

Lastly, we show that $a^m b^m = (ab)^m$ for all real $a, b \neq 0$ and $m \in \mathbb{Z}$.

Proof. We have the following cases:

Case: $m \in \mathbb{Z}_+$. The result then follows immediately from Exercise 4.6.

Case: $m = 0$. Then we have $a^m b^m = a^0 b^0 = 1 \cdot 1 = 1 = (ab)^0 = (ab)^m$ by the definition of a 0 exponent.

Case: $m \in \mathbb{Z}_-$. Then there is a $k \in \mathbb{Z}_+$ such that $m = -k$. Then we have

$$a^m b^m = a^{-k} b^{-k}$$

$$
\begin{aligned}
&= \frac{1}{a^k} \cdot \frac{1}{b^k} && \text{(by the definition of negative exponents)} \\
&= \left(\frac{1}{a}\right)^k \left(\frac{1}{b}\right)^k && \text{(by Lemma 4.7.2)} \\
&= \left(\frac{1}{a} \cdot \frac{1}{b}\right)^k && \text{(by Exercise 4.6 since } k \in \mathbb{Z}_+) \\
&= \left(\frac{1}{ab}\right)^k && \text{(by Exercise 4.1m)} \\
&= \frac{1}{(ab)^k} && \text{(by Lemma 4.7.2)} \\
&= (ab)^{-k} && \text{(by the definition of negative exponents)} \\
&= (ab)^m .
\end{aligned}
$$

Therefore in all cases the result has been shown. □

Exercise 4.8

(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 $0 < a < 1$, $\inf \{a^n \mid n \in \mathbb{Z}_+\} = 0$. [Hint: Let $h = (1-a)/a$, and show that $(1+h)^n \geq 1 + nh$.]

Solution:

(a)

Proof. Suppose that A is an arbitrary nonempty set of real number that is bounded below by a. Now let $B = \{-x \mid x \in A\}$ and $b = -a$. First, we claim that b is an upper bound of B. So consider any $y \in B$ so that $y = -x$ for some $x \in A$. Then $a \leq x$ since a a lower bound of A. It then follows from Exercise 4.2d that $y = -x \leq -a = b$. Since $y \in B$ was arbitrary, this shows that b is an upper bound of B.

Since B is clearly nonempty (since A is), we have that B has a least upper bound $d = \sup B$ since the reals have the least upper bound property. We claim that $c = -d$ is the greatest lower bound of A. So first consider any $x \in A$ so that $y = -x \in B$. Then we have $y \leq d$ since $d = \sup B$. Hence $c = -d \leq -y = x$ again by Exercise 4.2d. Since $x \in A$ was arbitrary, this shows that c is in fact a lower bound of A.

Now suppose that x is any lower bound of A. Then, by the same argument as above for $b = -a$, we have that $y = -x$ is an upper bound of B. It then follows that $d \leq y$ since d is the *least* upper bound of B. Then, again by Exercise 4.2d, we have $x = -(-x) = -y \leq -d = c$, which shows that c is in fact the greatest lower bound since x was arbitrary. This completes the proof. □

(b)

Proof. First, let $A = \{1/n \mid n \in \mathbb{Z}_+\}$ so that we must show that $\inf A = 0$. For any $x \in A$ we have that $x = 1/n$ for some $n \in \mathbb{Z}_+$. Then $n > 0$ so that $x = 1/n > 0$ also by Exercise 4.2i. Hence $0 \leq x$ is true, which shows that 0 is a lower bound of A since x was arbitrary.

Now consider any $x > 0$ so that also $1/x > 0$ by Exercise 4.2i. Then, by the Archimedean ordering property there is an $n \in \mathbb{Z}_+$ such that $n > 1/x > 0$ (since otherwise $1/x$ would be an upper bound of $\mathbb{Z}_+$). It then follows from Exercise 4.2j that $1/n < 1/(1/x) = x$. Since clearly $1/n \in A$ we have that x is *not* a lower bound of A. Since $x > 0$ was arbitrary, this shows that 0 is the greatest lower bound of A since, by the contrapositive, x being a lower bound of A implies that $x \leq 0$. $\square$

(c)

Proof. Consider any real a where $0 < a < 1$. First we show that the set $\{1/a^n \mid n \in \mathbb{Z}_+\}$ has no upper bound. To this end define $h = (1-a)/a = 1/a - 1$ so that $1 + h = 1 + (1/a - 1) = 1/a$. Clearly we have

$$\begin{aligned} a &< 1 \\ -a &> -1 \\ 1 - a &> 1 - 1 = 0 \\ \frac{1-a}{a} &> \frac{0}{a} = 0 && \text{(since } a > 0\text{)} \\ h &> 0 \end{aligned}$$

so that $1 + h > 1 > 0$ and

$$\begin{aligned} h &> 0 \\ h^2 &> h \cdot 0 = 0 && \text{(since } h > 0\text{)} \\ nh^2 &> n \cdot 0 = 0 \end{aligned}$$

for any $n \in \mathbb{Z}_+$ since $n > 0$.

We show by induction that $(1+h)^n \geq 1 + nh$ for all $n \in \mathbb{Z}_+$. For $n = 1$ we clearly have $(1+h)^n = (1+h)^1 = 1 + h \geq 1 + h = 1 + 1 \cdot h = 1 + nh$. Now, supposing that $(1+h)^n \geq 1 + nh$,we have

$$\begin{aligned} (1+h)^{n+1} &= (1+h)^n(1+h) \\ &\geq (1+nh)(1+h) && \text{(since } 1 + h > 0\text{)} \\ &= 1 + nh + h + nh^2 \\ &\geq 1 + nh + h && \text{(since } nh^2 > 0\text{)} \\ &= 1 + (n+1)h\,, \end{aligned}$$

which completes the induction. So consider any real x. Then, since we know that $\mathbb{Z}_+$ has no upper bound, there is an $n \in \mathbb{Z}_+$ where $n > x/h$ (noting that $h > 0$) so that

$$\begin{aligned} n &> x/h \\ nh &> (x/h)h = x && \text{(since } h > 0\text{)} \\ 1 + nh &> 1 + x > x\,. \end{aligned}$$

Then we have $1/a^n = (1/a)^n = (1+h)^n \geq 1 + nh > x$, which shows that the set $\{1/a^n \mid n \in \mathbb{Z}_+\}$ is unbounded above since x was arbitrary.

Now we show the main result. Let $A = \{a^n \mid n \in \mathbb{Z}_+\}$ so that we must show that $\inf A = 0$. First we show by induction that 0 is a lower bound of A. For $n = 1$ we clearly have $a^n = a^1 = a \geq 0$. Then, if $a^n \geq 0$, we have $a^{n+1} = a^n \cdot a \geq 0 \cdot a = 0$ since $a > 0$. This completes the induction so that clearly 0 is indeed a lower bound of A.

Now consider any real $x > 0$ so that $1/x > 0$ also. Then, by what was shown above, we know that there is an $n \in \mathbb{Z}_+$ such that $1/a^n > 1/x > 0$. We then have $a^n = 1/(1/a^n) < 1/(1/x) = x$ by

Exercise 4.2j. This shows that x is *not* a lower bound of A since obviously $a^n \in A$. It then follows that 0 is the *greatest* lower bound of A since $x > 0$ was arbitrary, because, by the contrapositive, x being a lower bound of A implies that $x \leq 0$. Hence $0 = \inf A$ as desired. □

Exercise 4.9

(a) Show that every nonempty subset of $\mathbb{Z}$ that is bounded above has a largest element.
(b) If $x \notin \mathbb{Z}$, show that 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:

Lemma 4.9.1. *The set of integers $\mathbb{Z}$ is an inductive set that has no lower or upper bounds in $\mathbb{R}$.*

Proof. First we show that $\mathbb{Z}$ is inductive. Clearly $1 \in \mathbb{Z}$ since $1 \in \mathbb{Z}_+ \subset \mathbb{Z}$. Now suppose that $n \in \mathbb{Z}$ so that clearly $n + 1 \in \mathbb{Z}$ by Exercise 4.5d since $1 \in \mathbb{Z}$.

Next, consider any $x \in \mathbb{R}$. Then we know that $\mathbb{Z}_+$ has no upper bound so that there is an $n \in \mathbb{Z}_+$ such that $n > x$, and clearly $n \in \mathbb{Z}$ since $\mathbb{Z}_+ \subset \mathbb{Z}$. By the same token there is an $m \in \mathbb{Z}_+$ such that $m > -x$. But then we have $-m < -(-x) = x$ by Exercise 4.2d, and $-m \in \mathbb{Z}_-$ so that also $-m \in \mathbb{Z}$ since $\mathbb{Z}_- \subset \mathbb{Z}$. Since x was arbitrary, this shows that $\mathbb{Z}$ is not bounded above or below. □

Lemma 4.9.2. *There is no integer n such that $0 < n < 1$.*

Proof. Suppose to the contrary that $n \in \mathbb{Z}$ and $0 < n < 1$. Let $S = \{k \in \mathbb{Z} \mid 0 < k < 1\}$ so that clearly $n \in S$ so that $S \neq \varnothing$. Also since $0 < k$ and $k \in \mathbb{Z}$ for any $k \in S$, clearly $S \subset \mathbb{Z}_+$. Thus S is a nonempty subset of positive integers so that it has a smallest element m by the well-ordering property. Since $m \in S$ we have $0 < m < 1$ and hence $m^2 = m \cdot m < 1 \cdot m = m < 1$ by property (6) since $m > 0$. By the same property clearly $0 = 0 \cdot m < m \cdot m = m^2$ as well so that $0 < m^2 < 1$. Also, clearly $m^2 = m \cdot m \in \mathbb{Z}$ by Exercise 4.5e since $m \in \mathbb{Z}$, and so $m^2 \in S$. However, this cannot be since m is the smallest element of S and yet $m^2 < m$. Therefore we have a contradiction, which proves the result. □

Corollary 4.9.3. *For any integer n, there is no integer a such that $n < a < n + 1$.*

Proof. Consider any $n \in \mathbb{Z}$ and suppose to the contrary that there is an $a \in \mathbb{Z}$ such that $n < a < n+1$. First, we have $n - a \in \mathbb{Z}$ by Exercise 4.5d since $a, n \in \mathbb{Z}$. Also, $n < a$ clearly implies that $0 < a - n$. Similarly, $a < n + 1$ means that $a - n < 1$. But then we have that $a - n$ is an integer where $0 < a - n < 1$, which contradicts Lemma 4.9.2. Thus it must be the case that there is no such integer a. □

Main Problem.

(a)

Proof. Suppose that A is a nonempty subset of $\mathbb{Z}$ and that it is bounded above by $\alpha \in \mathbb{R}$. Since $A \neq \varnothing$, there is an $a \in A$, so define $A' = \{n - a + 1 \mid n \in A\}$. First we claim that $\alpha' = \alpha - a + 1$ is an upper bound of A' So consider any $n' \in A'$ so that $n' = n - a + 1$ for some $n \in A$. Since α is an upper bound of A we have

$$n \leq \alpha$$

$$\begin{aligned} n - a &\le \alpha - a \\ n - a + 1 &\le \alpha - a + 1 \\ n' &\le \alpha' \,, \end{aligned}$$

which shows that α' is an upper bound of A' since n' was an arbitrary element. We also have that there is an $N' \in \mathbb{Z}_+$ such that $\alpha' < N'$ since $\mathbb{Z}_+$ has no upper bound.

Now let $B' = A' \cap \mathbb{Z}_+$. Then, for any $n' \in B'$, we have that $n' \in A'$ so that $n' \le \alpha' < N'$. Since also clearly $n' \in \mathbb{Z}_+$, we have that $n' \in S_{N'} = \{k \in \mathbb{Z}_+ \mid k < N'\} = \{1, \ldots, N' - 1\}$. Hence $B' \subset S_{N'}$ since n' was arbitrary. We also have that $1 \in A'$ since $a \in A$ and $a - a + 1 = 1$. Hence $1 \in B'$ since clearly also $1 \in \mathbb{Z}_+$ since it is inductive. Thus B' is a nonempty subset of $S_{N'}$ so that it has a largest element b' by Exercise 4.4a.

Since $b' \in B'$, we have that $b' \in A'$ so that there is a $b \in A$ such that $b' = b - a + 1$. We claim that b is the largest element of A. We already know that $b \in A$ so we need only show that it is also an upper bound of A. So consider any $n \in A$ so that clearly $n' = n - a + 1 \in A'$. Now, it follows from Exercise 4.5d that $n' \in \mathbb{Z}$ since $n, a, 1 \in \mathbb{Z}$. Thus we have the following:

Case: $n' \in \mathbb{Z}_+$. Then, clearly $n' \in A' \cap \mathbb{Z}_+ = B'$ so that $n' \le b'$ since b' is the largest element of B'.

Case: $n' \in \mathbb{Z}_- \cup \{0\}$. Then $n' < 1 \le b'$ since $1 \in B'$ and b' is the largest element of B'.

Thus in either case $n' \le b'$ is true so that

$$\begin{aligned} n' &\le b' \\ n - a + 1 &\le b - a + 1 \\ n - a &\le b - a \\ n &\le b \,, \end{aligned}$$

which shows that b is an upper bound and thus the largest element of A since n was arbitrary. □

(b)

Proof. Suppose an $x \in \mathbb{R}$ where $x \notin \mathbb{Z}$ and let $A = \{n \in \mathbb{Z} \mid n < x\}$. It follows from Lemma 4.9.1 that there is an $m \in \mathbb{Z}$ where $m < x$ since $\mathbb{Z}$ has no lower bounds. Hence by definition $m \in A$ so that $A \ne \varnothing$. Clearly also x is an upper bound of A so that A is a nonempty subset of $\mathbb{Z}$ that is bounded above. It then follows from part (a) that A has a largest element n, where clearly $n < x$ since $n \in A$.

Now, suppose for the moment that $n + 1 \le x$. Then, since $\mathbb{Z}$ is inductive (again by Lemma 4.9.1) and $n \in \mathbb{Z}$, we have that $n + 1 \in \mathbb{Z}$ as well. But $x \notin \mathbb{Z}$ so that it must be that $n + 1 \ne x$, and hence $n + 1 < x$. Then $n + 1 \in A$ so that $n + 1 \le n$ since n is the largest element of A. However, this contradicts the obvious fact that $n + 1 > n$ so that it must be that $n + 1 \le x$ is not true. Hence $n + 1 > x$ and thus we have shown that $n < x < n + 1$.

Lastly, suppose that there is an integer m such that $m < x < m + 1$. Then $m \in A$ so that $m \le n$ since n is the largest element of A. Suppose for a moment that $m < n$. Then we would have $m < n < x < m + 1$ so that n is an integer between m and $m + 1$, which violates Corollary 4.9.3. Thus is has to be that $m = n$ (since $m \le n$), which shows that n is the unique integer such that $n < x < n + 1$. □

(c)

Proof. Suppose that $x, y \in \mathbb{R}$ and $x - y > 1$. If $x \in \mathbb{Z}$ then let $n = x - 1$ so that clearly $n \in \mathbb{Z}$ by Exercise 4.5d. First, we have

$$x - y > 1$$

$$\begin{aligned} x &> 1 + y \\ x - 1 &> y \\ n &> y\,. \end{aligned}$$

We also clearly have $n = x - 1 < x$ so that $y < n < x$.

On the other hand, if $x \notin \mathbb{Z}$, then we know from part (b) that there is a unique integer n such that $n < x < n + 1$. We also have that

$$\begin{aligned} x &< n + 1 \\ 1 < x - y &< n + 1 - y \\ 0 &< n - y \\ y &< n \end{aligned}$$

so that again $y < n < x$.

Hence in both cases we have found an integer n such that $y < n < x$, which proves the result. □

(d)

Proof. Suppose that $x, y \in \mathbb{R}$ where $y < x$. Then $0 < x - y$ so that $1/(x - y)$ exists.. Since $\mathbb{Z}_+$ is unbounded above there is a $b \in \mathbb{Z}_+$ where $b > 1/(x - y)$. Hence

$$\begin{aligned} b &> \frac{1}{x - y} \\ b(x - y) &> 1 && \text{(since } x - y > 0\text{)} \\ bx - by &> 1\,. \end{aligned}$$

It then follows from part (c) that there is an integer a such that $by < a < bx$. We then have that $y < a/b < x$ since $b > 0$ (since $b \in \mathbb{Z}_+$). This shows the result since clearly a/b is rational because $a, b \in \mathbb{Z}$. □

Exercise 4.10

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

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

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

Lemma 4.10.1. *If $x \in \mathbb{R}$ and $x^2 < 1$, then $x < 1$ also.*

Proof. Suppose that $x \geq 1$. If $x = 1$ then clearly $x^2 = 1^1 = 1$. On the other hand, if $x > 1$ then clearly $x^2 = x \cdot x > 1 \cdot x = x > 1$ by property (6) since $x > 1 > 0$. Thus in either case $x^2 \geq 1$ so that we have shown that $x \geq 1$ implies that $x^2 \geq 1$. It then follows that $x^2 < 1$ implies $x < 1$ by the contrapositive. □

Lemma 4.10.2. *If* $0 < y < x$ *then* $0 < y^2 < x^2$.

Proof. Supposing that $0 < y < x$, we have $0 = 0 \cdot y < y \cdot y = y^2 = y \cdot y < x \cdot y = y \cdot x < x \cdot x = x^2$ all by property (6) since both x and y are positive. □

Main Problem.

(a)

Proof. First, we know that $0 \leq h < 1$. If $h = 0$ then clearly $h = 0 = 0^2 = h^2$ so that $0 \leq h^2 \leq h$ is true. If $h \neq 0$ then $0 < h < 1$ so that $0 = 0 \cdot h < h \cdot h = h^2 < 1 \cdot h = h$ by property (6) since $h > 0$ so that again $0 \leq h^2 \leq h$ is true.

We then have

$$\begin{aligned} (x+h)^2 &= (x+h)(x+h) \\ &= x^2 + 2xh + h^2 \\ &\leq x^2 + 2xh + h && \text{(since } h^2 \leq h\text{)} \\ &= x^2 + h(2x+1)\,. \end{aligned}$$

Also

$$\begin{aligned} (x-h)^2 &= (x-h)(x-h) \\ &= x^2 - 2xh + h^2 \\ &\geq x^2 - 2xh + 0 && \text{(since } h^2 \geq 0\text{)} \\ &= x^2 - h(2x)\,, \end{aligned}$$

which show the desired results. □

(b) We modify this result so that the h in the second part is not just positive but also $h < x$. In fact, without this stipulation, the theorem becomes obvious since any arbitrarily large h will suffice. Because then then $x - h$ is arbitrarily large in magnitude (but negative) so that $(x-h)^2$ can be made arbitrarily large so that of course $(x-h)^2 > a$. Adding the stipulation that $0 < h < x$ makes the theorem more useful and is necessary for it to be of use in part (c) below.

Proof. Suppose that $x > 0$. Then clearly $2x > 2 \cdot 0 = 0$ as well. Also it then follows that $2x + 1 > 1 > 0$.

If $x^2 < a$ then clearly $0 < a - x^2$. Hence we have that $0 < (a - x^2)/(2x+1)$ by Exercise 4.2 parts (i) and (h) since both $a - x^2$ and $2x + 1$ are positive. So let $y = \min(1, (a - x^2)/(2x+1))$ so that clearly both $y \leq 1$ and $y \leq (a - x^2)/(2x+1)$. Since $0 < 1$ and $0 < (a - x^2)/(2x+1)$, we have that $0 < y$ so that it follows from Exercise 4.9d that there is a rational h such that $0 < h < y$. Hence $0 < h < y \leq 1$ so that, by part (a), we have

$$\begin{aligned} (x+h)^2 &\leq x^2 + h(2x+1) \\ &< x^2 + \left(\frac{a - x^2}{2x+1}\right)(2x+1) && \text{(since } h < y \leq (a-x^2)/(2x+1) \text{ and } 2x+1 > 0\text{)} \\ &= x^2 + (a - x^2) \end{aligned}$$

$$= a\,.$$

If $x^2 > a$ then clearly $x^2 - a > 0$. Then we have again that $(x^2-a)/(2x)$ is positive since we showed previously that $2x$ is. So let $y = \min(1, (x^2-a)/(2x), x)$ so that clearly $y \leq 1$, $y \leq (x^2-a)/(2x)$, and $y \leq x$. Since both 1, $(x^2-a)/(2x)$, and x are all positive it follows that $0 < y$ so that there is a rational h such that $0 < h < y$ by Exercise 4.9d. Therefore $0 > -h > -y$. Since $0 < h < y \leq 1$ we have by part (a) that

$$\begin{aligned}
(x-h)^2 &\geq x^2 - h(2x) \\
&> x^2 - \left(\frac{x^2-a}{2x}\right)(2x) && \text{(since } -h > -y \geq -(x^2-a)/(2x) \text{ and } 2x > 0) \\
&= x^2 - (x^2 - a) \\
&= a\,,
\end{aligned}$$

which show the desired results since clearly $0 < h < y \leq x$. □

(c)

Proof. Suppose that $a > 0$ and let $B = \{x \in \mathbb{R} \mid x^2 < a\}$.

If $a < 1$ then $0 < a < 1$ so that $a^2 = a \cdot a < 1 \cdot a = a$ so that a itself is in B (and of course a is positive). Now consider any $x \in B$ so that $x^2 < a$. Then $x^2 < a < 1$ so that also $x < 1$ by Lemma 4.10.1. Since $x \in B$ was arbitrary, this shows that 1 is an upper bound of B.

If $a \geq 1$ then $(1/2)^2 = 1/2^2 = 1/4 < 1 \leq a$ so that $1/2 \in B$ (and of course $1/2$ is positive). Now consider any $x \in B$ so that $x^2 < a$. If $x \leq 1$ then $x \leq 1 \leq a$. On the other hand, if $x > 1$ then $x^2 = x \cdot x > 1 \cdot x = x$ since $x > 1 > 0$ so that $x < x^2 < a$. Thus in both cases $x \leq a$ so that a is an upper bound of B since x was arbitrary.

Therefore in each case B contains a positive element (so that $b \neq \varnothing$) and B is bounded above. It then follows that B has a least upper bound b (so that $b = \sup B$). Clearly since B has a positive element x, it follows that $0 < x \leq b$ so that b is positive.

Now suppose that $b^2 < a$. Then by definition $b \in B$ so that b has to be the largest element of b since it is the least upper bound. Since we know that b is positive and $b^2 < a$, it follows from part (b) that there is an $h > 0$ where $(b+h)^2 < a$ and hence $b + h \in B$. However, since $h > 0$, it follows that $b < b + h$, which contradicts the fact that b is the greatest element of B. Hence it cannot be that $b^2 < a$.

So suppose that $b^2 > a$. Then again by part (b) there is an h where $0 < h < b$ such that $(b-h)^2 > a$. Now, since $h > 0$, it follows that $b - h < b$ so that $n - h$ is not an upper bound of B (since then b would not be the least upper bound). Hence there is an $x \in B$ such that $b - h < x$, noting that $x^2 < a$ by the definition of B. Since $h < b$, we have that $0 < b - h < x$ so that $(b-h)^2 < x^2 < a$ by Lemma 4.10.2. But this contradicts the established fact that $(b-h)^2 > a$ so that it cannot be that $b^2 > a$.

Thus the only possibility remaining is that $b^2 = a$ as desired. □

(d)

Proof. Suppose that b and c are positive and that $b^2 = c^2$. If it were the case that $b < c$ then $0 < b < c$ so that $0 < b^2 < c^2$ by Lemma 4.10.2 so that clearly $b^2 \neq c^2$. As this is a contradiction, it has to be that $b \geq c$. An analogous argument shows that $b > c$ also leads to a contradiction so that $b \leq c$. Hence it must be that $b = c$ as desired. □

Exercise 4.11

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

(a) Show that if m is odd, $m = 2n + 1$ for some $n \in \mathbb{Z}$. [Hint: Choose n so that $n < m/2 < n + 1$.]

(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 n and m are even. [Hint: Let n be the smallest element of the set $\{x \mid x \in \mathbb{Z}_+ \text{ and } x \cdot a \in \mathbb{Z}_+\}$.]

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

Solution:

Lemma 4.11.1. *If* $n, m \in \mathbb{Z}$ *and* $n < m$*, then* $n + 1 \le m$ *and* $n \le m - 1$.

Proof. Suppose that $n + 1 > m$ so that $n < m < n + 1$, which violates Corollary 4.9.3 since $m \in \mathbb{Z}$. Thus it has to be that $n + 1 \le m$. From this it immediately follows that $n = n + 1 - 1 \le m - 1$ by simply subtracting 1 from both sides of the previous inequality. □

Lemma 4.11.2. *An integer* m *is even if and only if* $m = 2n$ *for some integer* n.

Proof. (⇒) Supposing that m is even, then $n = m/2 \in \mathbb{Z}$. Then clearly $m = 2n$.

(⇐) Now suppose that $m = 2n$ for some integer n. Then clearly $m/2 = n$ is an integer so that m is even by definition. □

Lemma 4.11.3. *An integer* a *is odd if and only if* a^2 *is also odd.*

Proof. (⇒) Suppose that a is odd so that $a = 2n + 1$ for some integer n (this is shown in part (a) below, which does not depend on this lemma). Then

$$a^2 = a \cdot a = (2n + 1)(2n + 1) = 4n^2 + 2n + 2n + 1 = 4n^2 + 4n + 1 = 2\left[2(n^2 + n)\right] + 1$$

noting that clearly $2(n^2 + n)$ is an integer since n is. Hence a^2 is odd again by what will be shown in part (a).

(⇐) We prove this by contrapositive, so suppose that a is not odd so that it must be even. Therefore $a = 2n$ for some integer n by Lemma 4.11.2. Then $a^2 = a \cdot a = (2n)(2n) = 4n^2 = 2(2n^2)$ so that a^2 is even since clearly $2n^2$ is an integer since n is. Thus a^2 is not odd. □

Main Problem.

(a) Here we show the converse as well, i.e. we show that m is odd if and only if $m = 2n + 1$ for some $n \in \mathbb{Z}$.

Proof. (⇒) Suppose that m is odd so that by definition $m/2 \notin \mathbb{Z}$. It then follows from Exercise 4.9b that there is a unique integer n such that $n < m/2 < n + 1$. We then have that $2n < m < 2(n + 1) = 2n + 2$ since obviously $2 > 0$. Hence by Lemma 4.11.1 we have that $2n + 1 \le m$ and also $m \le 2n + 2 - 1 = 2n + 1$. Therefore it has to be that $m = 2n + 1$ as desired.

(⇐) Now suppose that there is an $n \in \mathbb{Z}$ such that $m = 2n + 1$. Then we have that

$$\frac{m}{2} = \frac{2n + 1}{2} = n + \frac{1}{2}\,.$$

We then clearly have that $n = n + 0 < n + 1/2 < n + 1$ since $0 < 1/2 < 1$ so that $m/2 = n + 1/2$ cannot be an integer by Corollary 4.9.3. Hence m is odd by definition. □

(b)

Proof. Suppose that p and q are odd so that $p = 2k + 1$ and $q = 2m + 1$ for some $k, m \in \mathbb{Z}$ by part (a). We then have that

$$p \cdot q = (2k + 1)(2m + 1) = 4km + 2m + 2k + 1 = 2(2km + m + k) + 1$$

so that $p \cdot q$ is odd by what was shown in part (a) since clearly $2km + m + k \in \mathbb{Z}$ by Exercise 4.5 since k and m are integers.

Now we show by induction on n that p^n is odd for any $n \in \mathbb{Z}_+$. First, for $n = 1$ we clearly have $p^n = p^1 = p$ is odd by supposition. Then, if we assume that p^n is odd, we have that the product $p^{n+1} = p^n \cdot p$ is odd as well by what was just shown since both p^n and p are odd. This completes the induction. □

(c)

Proof. Suppose that $a > 0$ is rational. Then $a = p/q$ for some integers p and q. Clearly it cannot be that $q = 0$, and if $q < 0$ then $q = -b$ for some $b \in \mathbb{Z}_+$. Then we have $a = p/q = p/(-b) = (-p)/b$ so that $ab = -p$. Furthermore, since a and b are both positive, we have that $ab = -p$ is positive by Exercise 4.2h. Thus clearly $-p \in \mathbb{Z}_+$ since $p \in \mathbb{Z}$.

Now, let $X = \{x \in \mathbb{Z}_+ \mid ax \in \mathbb{Z}_+\}$. Since we just showed that $b \in \mathbb{Z}_+$ and $ab = -p \in \mathbb{Z}_+$ it follows that $b \in X$. Since clearly $X \subset \mathbb{Z}_+$ and X is nonempty (since $b \in X$), it has a smallest element n by the well-ordering property. Letting $m = an$, we clearly have that $m \in \mathbb{Z}_+$ since $n \in X$. Then, we have $a = m/n$, noting again that $m, n \in \mathbb{Z}_+$.

To show that not both m and n are even, suppose to the contrary that they *are* both even. Then by Lemma 4.11.2 we have that $m = 2k$ and $n = 2l$ for some $k, l \in \mathbb{Z}$. Clearly then $k = m/2$ and $l = n/2$ so that both k and l are positive by Exercise 4.2h since m and n (and $1/2$) are. Hence $k, l \in \mathbb{Z}_+$. We have $a = m/n = 2k/2l = k/l$ so that $al = k$, which implies that $l \in X$ since l and $al = k$ are both in $\mathbb{Z}_+$. However, we also have that $l = n/2 < n$ since $n > 0$, which contradicts the fact that n is the smallest element of X. Thus it has to be the case that not both m and n are even. □

(d) This is one of the most famous proofs in all of mathematics, and is often used as an example of mathematical proofs since it can be understood by most laymen.

Proof. Obviously we take $\sqrt{2}$ to be the unique positive real number such that $(\sqrt{2})^2 = 2$ as was shown to exist in Exercise 4.10. Suppose to the contrary that $\sqrt{2}$ is rational so that $\sqrt{2} = a/b$ for $a, b \in \mathbb{Z}_+$ where not both a and b are even by part (c) since $\sqrt{2} > 0$. We therefore have that $2 = (\sqrt{2})^2 = (a/b)^2 = a^2/b^2$ so that $2b^2 = a^2$. Since b^2 is an integer (clearly, since b is and $b^2 = b \cdot b$) it follows from Lemma 4.11.2 that a^2 is even. This means that a itself is even by Lemma 4.11.3. Hence $a = 2n$ for some integer n so that $a^2 = (2n)^2 = 4n^2$. From before, we then have $2b^2 = a^2 = 4n^2$ so that clearly $b^2 = 2n^2$, from which it follows as before that b^2 and therefore b itself is even by Lemmas 4.11.2 and 4.11.3. However, this is a contradiction since we previously established that a and b cannot both be even! So it has to be that $\sqrt{2}$ is not rational and is therefore irrational as desired. □

§5 Cartesian Products

Exercise 5.1

Show that there is a bijective correspondence of $A \times B$ with $B \times A$.

Solution:

Proof. We define a function $f : A \times B \to B \times A$. For any element $(a, b) \in A \times B$ we set $f(a, b) = (b, a)$, noting that of course $a \in A$ and $b \in B$. It should be obvious then that $f(a, b) = (b, a) \in B \times A$ so that $B \times A$ can be the range of f.

First we show that f is injective. To this end consider (a_1, b_1) and (a_2, b_2) in $A \times B$ where $(a_1, b_1) \neq (a_2, b_2)$. Of course we have that $f(a_1, b_1) = (b_1, a_1)$ and $f(a_2, b_2) = (b_2, a_2)$. Since $(a_1, b_1) \neq (a_2, b_2)$ clearly either $a_1 \neq a_2$ or $b_1 \neq b_2$. In either case it should be clear that $f(a_1, b_1) = (b_1, a_1) \neq (b_2, a_2) = f(a_2, b_2)$, which shows that f is injective since (a_1, b_1) and (a_2, b_2) were arbitrary.

It is very easy to that f is also surjective since, for any $(b, a) \in B \times A$, clearly $(a, b) \in A \times B$ and $f(a, b) = (b, a)$. Hence f is a bijection as desired. Note that if $A \times B = \varnothing$ then $f = \varnothing$ as well, which is vacuously a bijective function since it must be that $B \times A = \varnothing$ as well (because either $A = \varnothing$ or $B = \varnothing$). □

Exercise 5.2

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

$$A_1 \times \cdots \times A_n \qquad \text{with} \qquad (A_1 \times \cdots \times A_{n-1}) \times A_n \,.$$

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

Solution:

Lemma 5.2.1. *If $n \in \mathbb{Z}_+$ is even, then $n/2 \in \mathbb{Z}_+$. If $n \in \mathbb{Z}_+$ is odd, then $(n + 1)/2 \in \mathbb{Z}_+$.*

Proof. First, suppose that $n \in \mathbb{Z}_+$ is even. Then by definition $n/2$ is an integer. However, since both n and $1/2$ are positive, it follows from Exercise 4.2h that $n \cdot (1/2) = n/2$ is positive also so that $n/2 \in \mathbb{Z}_+$.

Now, suppose that $n \in \mathbb{Z}_+$ is odd so that $n = 2k + 1$ for some integer k by Exercise 4.11a. Then

$$\frac{n+1}{2} = \frac{(2k+1)+1}{2} = \frac{2k+2}{2} = \frac{2(k+1)}{2} = k + 1\,,$$

which is clearly an integer since k is. Moreover, we have $n + 1 > n > 0$ since $n \in \mathbb{Z}_+$ and again $1/2 > 0$ so that $(n + 1) \cdot (1/2) = (n + 1)/2$ is positive by Exercise 4.2h. Thus $(n + 1)/2 \in \mathbb{Z}_+$. □

Main Problem.

(a)

Proof. For brevity, let $X = A_1 \times \cdots \times A_n$ and $Y = (A_1 \times \cdots \times A_{n-1}) \times A_n$. Suppose that $n > 1$ so that X and Y make sense. We construct a bijective function $f : X \to Y$. For any $\mathbf{x} = (x_1, \ldots x_n) \in X$ we have that $x_i \in A_i$ for $1 \leq i \leq n$. So set $f(x) = ((x_1, \ldots, x_{n-1}), x_n)$, which is clearly an element of Y.

To see that f is injective consider $\mathbf{x} = (x_1, \ldots, x_n)$ and $\mathbf{y} = (y_1, \ldots, y_n)$ in X where $\mathbf{x} \neq \mathbf{y}$. It then follows that there must be an $i \in \{1, \ldots, n\}$ where $x_i \neq y_i$. Let $\mathbf{x}' = (x_1, \ldots, x_{n-1})$ and $\mathbf{y}' = (y_1, \ldots, y_{n-1})$ so that clearly $f(\mathbf{x}) = (\mathbf{x}', x_n)$ and $f(\mathbf{y}) = (\mathbf{y}', y_n)$. Now, if $i = n$, then clearly $f(\mathbf{x}) = (\mathbf{x}', x_n) \neq (\mathbf{y}', y_n) = f(\mathbf{y})$ since $x_n = x_i \neq y_i = y_n$. On the other hand, if $i \neq n$ then it has to be that $i < n$, and hence $i \leq n-1$. It then follows that $\mathbf{x}' = (x_1, \ldots, x_{n-1}) \neq (y_1, \ldots y_{n-1}) = \mathbf{y}'$ so that then $f(\mathbf{x}) = (\mathbf{x}', x_n) \neq (\mathbf{y}', y_n) = f(\mathbf{y})$ again. Since $\mathbf{x}$ and $\mathbf{y}$ were arbitrary, this shows that f is indeed injective.

Now consider any $\mathbf{y} = ((x_1, \ldots, x_{n-1}), x_n) \in Y$ and let $\mathbf{x} = (x_1, \ldots x_n)$. It should be obvious that both $\mathbf{x} \in X$ and $f(\mathbf{x}) = \mathbf{y}$ so that f is surjective. Hence f is a bijective function as desired. □

(b)

Proof. First let $A = A_1 \times A_2 \times \cdots$ and $B = B_1 \times B_2 \times \cdots$. We construct a bijective $f : A \to B$. So, for any $\mathbf{a} \in A$, we have that $\mathbf{a} = (a_1, a_2, \ldots)$, where $a_i \in A_i$ for any $i \in \mathbb{Z}_+$. Then, for any $i \in \mathbb{Z}_+$, define $b_i = (a_{2i-1}, a_{2i})$ so that clearly $b_i \in A_{2i-1} \times A_{2i} = B_i$. We then have that $\mathbf{b} = (b_1, b_2, \ldots) \in B_1 \times B_2 \times \cdots = B$. So set $f(\mathbf{a}) = \mathbf{b}$ so that f is a function from A to B.

To show that f is injective, consider $\mathbf{a} = (a_1, a_2, \ldots)$ and $\mathbf{a}' = (a_1', a_2', \ldots)$ in A where $\mathbf{a} \neq \mathbf{a}'$. For each $i \in \mathbb{Z}_+$, define $b_i = (a_{2i-1}, a_{2i})$ and $b_i' = (a_{2i-1}', a_{2i}')$ as above and set $\mathbf{b} = (b_1, b_2, \ldots)$ and $\mathbf{b}' = (b_1', b_2', \ldots)$ so that clearly $f(\mathbf{a}) = \mathbf{b}$ and $f(\mathbf{a}') = \mathbf{b}'$. Since $\mathbf{a} \neq \mathbf{a}'$, it follows that there must be an $i \in \mathbb{Z}_+$ where $a_i \neq a_i'$.

Case: i is even. Then let $j = i/2$ so that $j \in \mathbb{Z}_+$ by Lemma 5.2.1. We also clearly have that $i = 2j$ so that $b_j = (a_{2j-1}, a_{2j}) \neq (a_{2j-1}', a_{2j}') = b_j'$ since $a_{2j} = a_i \neq a_i' = a_{2j}'$.

Case: i is odd. Then let $j = (i+1)/2$ so that $j \in \mathbb{Z}_+$ by Lemma 5.2.1. We then clearly have that $i = 2j-1$ so that $b_j = (a_{2j-1}, a_{2j}) \neq (a_{2j-1}', a_{2j}') = b_j'$ since $a_{2j-1} = a_i \neq a_i' = a_{2j-1}'$.

Hence in all cases we have that there is a $j \in \mathbb{Z}_+$ where $b_j \neq b_j'$. It then follows that $f(\mathbf{a}) = \mathbf{b} = (b_1, b_2, \ldots) \neq (b_1', b_2', \ldots) = \mathbf{b}' = f(\mathbf{a}')$ so that f is injective since $\mathbf{a}$ and $\mathbf{a}'$ were arbitrary.

Lastly, to show that f is surjective, consider any $\mathbf{b} \in B$ so that $\mathbf{b} = (b_1, b_2, \ldots)$ where $b_i \in B_i = A_{2i-1} \times A_{2i}$ for every $i \in \mathbb{Z}_+$. Then, for any $i \in \mathbb{Z}_+$, $b_i = (a_i', a_i'')$ where $a_i' \in A_{2i-1}$ and $a_i'' \in A_{2i}$. So consider any $j \in \mathbb{Z}_+$. If j is even, then $i = j/2 \in \mathbb{Z}_+$ by Lemma 5.2.1. Clearly also $j = 2i$. So, define $a_j = a_i''$ so that $a_j = a_{2i} = a_i'' \in A_{2i} = A_j$. On the other hand, if j is odd, then $i = (j+1)/2 \in \mathbb{Z}_+$ again by Lemma 5.2.1. Then clearly $j = 2i - 1$. So, here let $a_j = a_i'$ so that $a_j = a_{2i-1} = a_i' \in A_{2i-1} = A_j$. Hence $a_j \in A_j$ for all $j \in \mathbb{Z}_+$ so that $\mathbf{a} = (a_1, a_2, \ldots) \in A$. Then, for any $i \in \mathbb{Z}_+$, we have $b_i = (a_i', a_i'') = (a_{2i-1}, a_{2i}) \in A_{2i-1} \times A_{2i} = B_i$ so that by definition $f(\mathbf{a}) = \mathbf{b} = (b_1, b_2, \ldots)$. This shows that f is surjective since $\mathbf{b}$ was arbitrary.

This completes the proof that f is bijective so that the desired result follows. □

Exercise 5.3

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 products)

(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? (We will return to this question in the exercises of §19.)

(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:

(a)

Proof. Suppose that $\mathbf{b} \in B$ so that $\mathbf{b} = (b_1, b_2, \ldots)$ where $b_i \in B_i$ for every $i \in \mathbb{Z}_+$. Consider any such $i \in \mathbb{Z}_+$ so that $b_i \in B_i$. Then also $b_i \in A_i$ since $B_i \subset A_i$. Since i was arbitrary, $b_i \in A_i$ for every $i \in \mathbb{Z}_+$ so that $\mathbf{b} = (b_1, b_2, \ldots) \in A_1 \times A_2 \times \cdots = A$. Since $\mathbf{b}$ was arbitrary, this shows that $B \subset A$. Note that we ignore the function range technicality issue mentioned above. $\square$

(b)

Proof. Suppose that $B \subset A$. Since $B \neq \varnothing$, there is a $\mathbf{b}' \in B$ so that $\mathbf{b}' = (b'_1, b'_2, \ldots)$ where $b'_i \in B_i$ for every $i \in \mathbb{Z}_+$. Now consider any $i \in \mathbb{Z}_+$ and $b_0 \in B_i$. Then define

$$b_j = \begin{cases} b_0 & j = i \\ b'_j & j \neq i \end{cases}$$

for any $j \in \mathbb{Z}_+$. Clearly we have that $b_j \in B_j$ for any $j \in \mathbb{Z}_+$ so that $\mathbf{b} = (b_1, b_2, \ldots) \in B_1 \times B_2 \times \cdots = B$. It then follows that also $\mathbf{b} \in A$ since $B \subset A$. Hence $b_j \in A_j$ for every $j \in \mathbb{Z}_+$. In particular, we have $b_0 = b_i \in A_i$. Since b_0 was arbitrary, this shows that $B_i \subset A_i$, and since i was arbitrary, this shows the desired result. $\square$

(c)

Proof. Suppose that A is nonempty so that there is an $\mathbf{a} \in A$. Then, since $A = A_1 \times A_2 \times \cdots$, it follows that $\mathbf{a} = (a_1, a_2, \ldots)$ where $a_i \in A_i$ for every $i \in \mathbb{Z}_+$. Therefore, for any such $i \in \mathbb{Z}_+$, we have that $a_i \in A_i$ so that $A_i \neq \varnothing$. Hence every A_i is nonempty as desired since i was arbitrary. $\square$

Consider the converse. Suppose that each A_i is nonempty (for $i \in \mathbb{Z}_+$). Then there is an $a_i \in A_i$ for every $i \in \mathbb{Z}_+$ so that $\mathbf{a} = (a_1, a_2, \ldots) \in A_1 \times A_2 \times \cdots = A$ so that then $A \neq \varnothing$. While this may seem like an innocuous argument, especially out of the context of axiomatic set theory, it actually requires the Axiom of Choice. The reason is that, in the general case when each A_i may have more than one element, or even an infinite number of elements, we have to choose a specific a_i in each A_i. Since the index set $\mathbb{Z}_+$ is infinite, an infinite number of these choices must be made, which is precisely when the Axiom of Choice is required. If the index set was finite, then the axiom would not be needed.

(d) First, let $C_i = A_i \cup B_i$ for every $i \in \mathbb{Z}_+$, and let $C = C_1 \times C_2 \times \cdots$, so that we are asked to compare C with $A \cup B$.

We claim that $A \cup B \subset C$ but that C is *not* generally a subset of $A \cup B$.

Proof. First consider any $\mathbf{x} \in A \cup B$ so that $\mathbf{x} \in A$ or $x \in B$. If $\mathbf{x} \in A$ then it has to be that $\mathbf{x} = (x_1, x_2, \ldots)$ where $x_i \in A_i$ for every $i \in \mathbb{Z}_+$. Consider then any such $i \in \mathbb{Z}_+$. Then $x_i \in A_i$ so that clearly $x_i \in A_i \cup B_i = C_i$. Since i was arbitrary, we conclude that $\mathbf{x} = (x_1, x_2, \ldots) \in C_1 \times C_2 \times \cdots = C$. An analogous argument shows that $\mathbf{x} \in C$ when $\mathbf{x} \in B$ as well. Hence $A \cup B \subset C$ since $\mathbf{x}$ was arbitrary.

To show that C is *not* a subset of $A \cup B$ in general, consider the following counterexample. Let $A_1 = \varnothing$ and $A_i = \{1\}$ for every $i \in \mathbb{Z}_+$ where $i > 1$. Also let $B_i = \{2\}$ for every $i \in \mathbb{Z}_+$. Now, it follows from the contrapositive of part (c) that $A = \varnothing$ since $A_1 = \varnothing$. We also clearly have $B = B_1 \times B_2 \times \cdots = \{(2, 2, \ldots)\}$ so that $A \cup B = \varnothing \cup B = B = \{(2, 2, \ldots)\}$. Clearly $C_1 = A_1 \cup B_1 = \varnothing \cup \{2\} = \{2\}$ while, for $i > 1$ we have $C_i = A_i \cup B_i = \{1\} \cup \{2\} = \{1, 2\}$. It then follows that, for $a_1 = 2$ and $a_i = 1$ for $i > 1$, we have $\mathbf{a} = (a_1, a_2, \ldots) = (2, 1, 1, \ldots) \in C_1 \times C_2 \times \cdots = C$. However, clearly $\mathbf{a} \notin A \cup B$, which suffices to show that C cannot be a subset of $A \cup B$ in general. $\square$

Now let $C_i = A_i \cap B_i$ for every $i \in \mathbb{Z}_+$ so that we are asked to compare $C = C_1 \times C_2 \times \cdots$ and $A \cap B$.

Here we claim that in fact $A \cap B = C$.

Proof. First consider any $\mathbf{x} \in A \cap B$ so that $\mathbf{x} \in A$ and $\mathbf{x} \in B$. It then follows that $\mathbf{x} = (x_1, x_2, \ldots)$ where $x_i \in A_i$ for every $i \in \mathbb{Z}_+$ and $x_i \in B_i$ for every $i \in \mathbb{Z}_+$. Then, for any such $i \in \mathbb{Z}_+$, clearly $x_i \in A_i$ and $x_i \in B_i$ so that $x_i \in A_i \cap B_i = C_i$. We then have that $\mathbf{x} = (x_1, x_2, \ldots) \in C_1 \times C_2 \times \cdots =$

Exercise 5.4

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

(a) If $m \leq n$, 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 an injective 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 $l : X^\omega \times X^\omega \to X^\omega$.

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

NOTE: For part (f), older printings of the text say, "If $A \subset B$, find an injective map $m : X^A \to X^B$." This is assumed to be an error since the meaning of X^A and X^B are not defined in the text (though, for example, X^A would typically mean the set of functions from A to X) as well as the fact that it was changed.

Solution:

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

Proof. Suppose that $m \leq n$. Since $X \neq \varnothing$, there is an $x_0 \in X$. Now, for any $\mathbf{x} \in X^m$ we have that $\mathbf{x} = (x_1, \ldots, x_m)$ where each $x_i \in X$. Then define

$$y_i = \begin{cases} x_i & 1 \leq i \leq m \\ x_0 & m < i \leq n \end{cases}$$

for $i \in \{1, \ldots, n\}$. Clearly $y_i \in X$ for every $i \in \{1, \ldots, n\}$ so that $\mathbf{y} = (y_1, \ldots, y_n) \in X^n$. Then set $f(\mathbf{x}) = \mathbf{y}$ so that $f : X^m \to X^n$.

To show that f is injective consider $\mathbf{x}$ and $\mathbf{x}'$ in X^m so that $\mathbf{x} = (x_1, \ldots, x_m)$ and $\mathbf{x}' = (x'_1, \ldots, x'_m)$ where both x_i and x'_i are of course in X for any $i \in \{1, \ldots, m\}$. Also suppose that $\mathbf{x} \neq \mathbf{x}'$ so that it follows that there is an $i \in \{1, \ldots, m\}$ where $x_i \neq x'_i$. Let $\mathbf{y} = (y_1, \ldots, y_n) = f(\mathbf{x})$ and $\mathbf{y}' = (y'_1, \ldots, y'_n) = f(\mathbf{x}')$. Then, since clearly $1 \leq i \leq m$, we have $y_i = x_i \neq x'_i = y'_i$ by the definition of f. Hence we have $f(\mathbf{x}) = \mathbf{y} \neq \mathbf{y}' = f(\mathbf{x}')$, which shows that f is injective since $\mathbf{x}$ and $\mathbf{x}'$ were arbitrary. □

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

Proof. Consider any $\mathbf{x} \in X^m \times X^n$ so that $\mathbf{x} = (\mathbf{a}, \mathbf{b})$ where $\mathbf{a} \in X^m$ and $\mathbf{b} \in X^n$. Then we have that $\mathbf{a} = (a_1, \ldots, a_m)$ and $\mathbf{b} = (b_1, \ldots, b_n)$ where $a_i, b_j \in X$ for every $i \in \{1, \ldots, m\}$ and $j \in \{1, \ldots, n\}$. Then define

$$y_k = \begin{cases} a_k & 1 \leq k \leq m \\ b_{k-m} & m < k \leq m+n \end{cases}$$

for any $k \in \{1, \ldots, m+n\}$, noting that for $m < k \leq m+n$ we have $m+1 \leq k \leq m+n$, and hence $1 \leq k-m \leq n$ so that b_{k-m} is defined. Now set $g(\mathbf{x}) = \mathbf{y} = (y_1, \ldots, y_{m+n})$ so that clearly $g(\mathbf{x}) \in X^{m+n}$ since each $y_k \in X$. Thus g is a function from $X^m \times X^n$ to X^{m+n}.

To show that g is injective, consider any $\mathbf{x} = (\mathbf{a}, \mathbf{b})$ and $\mathbf{x}' = (\mathbf{a}', \mathbf{b}')$ in $X^m \times X^n$ where $\mathbf{x} \neq \mathbf{x}'$. Also let $\mathbf{y} = (y_1, \ldots, y_{m+n}) = g(\mathbf{x})$ and $\mathbf{y}' = (y'_1, \ldots, y'_{m+n}) = g(\mathbf{x}')$. Since $\mathbf{x} \neq \mathbf{x}'$, it must be that $\mathbf{a} \neq \mathbf{a}'$ or $\mathbf{b} \neq \mathbf{b}'$. In the former case we have that $\mathbf{a} = (a_1, \ldots, a_m)$ and $\mathbf{a}' = (a'_1, \ldots, a'_m)$ since they are both in X^m. Since $\mathbf{a} \neq \mathbf{a}'$ there is an $i \in \{1, \ldots, m\}$ where $a_i \neq a'_i$. Then, since clearly $1 \leq i \leq m$, we have that $y_i = a_i \neq a'_i = y'_i$. In the latter case we have that $\mathbf{b} = (b_1, \ldots, b_n)$ and $\mathbf{b}' = (b'_1, \ldots, b'_n)$ since they are both in X^n. Then, since $\mathbf{b} \neq \mathbf{b}'$, we have that there is an $i \in \{1, \ldots, n\}$ such that $b_i \neq b'_i$. Let $k = m + i$ so that clearly $k - m = i$. Also $m < m + i = k \leq m + n$ since $0 < 1 \leq i \leq n$ so that $y_k = b_{k-m} = b_i \neq b'_i = b'_{k-m} = y'_k$. Hence in both cases there is a $k \in \{1, \ldots, m+n\}$ such that $y_k \neq y'_k$ so that $g(\mathbf{x}) = \mathbf{y} = (y_1, \ldots, y_{m+n}) \neq (y'_1, \ldots, y'_{m+n}) = \mathbf{y}' = g(\mathbf{x}')$. Since $\mathbf{x}$ and $\mathbf{x}'$ were arbitrary, this shows that g is indeed injective.

Now consider any $\mathbf{y} = (y_1, \ldots, y_{m+n}) \in X^{m+n}$, and define $a_i = y_i$ for any $i \in \{1, \ldots, m\}$ and $b_j = y_{m+j}$ for any $j \in \{1, \ldots, n\}$, noting that y_{m+j} is defined since $0 < 1 \leq j \leq n$ implies that $m < m + j \leq m + n$. Then let $\mathbf{a} = (a_1, \ldots, a_m)$, $\mathbf{b} = (b_1, \ldots, b_n)$, and $\mathbf{x} = (\mathbf{a}, \mathbf{b})$ so that clearly $\mathbf{x} \in X^m \times X^n$. Let $\mathbf{y}' = g(\mathbf{x})$ as defined above so that $\mathbf{y}' = (y'_1, \ldots, y'_{m+n})$. Consider any $k \in \{1, \ldots, m+n\}$. If $1 \leq k \leq m$ then we have by the definition of g that $y'_k = a_k = y_k$. On the other hand, if $m < k \leq m + n$, then we have $y'_k = b_{k-m} = y_{m+(k-m)} = y_k$. Thus in both cases $y'_k = y_k$ so that clearly $g(\mathbf{x}) = \mathbf{y}' = (y'_1, \ldots, y'_{m+n}) = (y_1, \ldots, y_{m+n}) = \mathbf{y}$ since k was arbitrary. This shows that g is surjective since $\mathbf{y}$ was arbitrary.

Therefore we have shown that g is bijective as desired. □

(c) Find an injective map $h : X^n \to X^\omega$.

Proof. First, we know that $X \neq \varnothing$ so that there is an $x_0 \in X$. So, for any $\mathbf{x} = (x_1, \ldots, x_n) \in X^n$, define

$$y_i = \begin{cases} x_i & 1 \leq i \leq n \\ x_0 & n < i \end{cases}$$

for any $i \in \mathbb{Z}_+$. Then set $h(\mathbf{x}) = \mathbf{y} = (y_1, y_2, \ldots)$ so that clearly $h(\mathbf{x}) \in X^\omega$. Thus h is a function that maps X^n into X^ω.

To show that h is injective, consider $\mathbf{x}$ and $\mathbf{x}'$ in X^n where $\mathbf{x} \neq \mathbf{x}'$. Clearly we have that $\mathbf{x} = (x_1, \ldots, x_n)$ and $\mathbf{x}' = (x'_1, \ldots, x'_n)$, and let $\mathbf{y} = (y_1, y_2, \ldots) = h(\mathbf{x})$ and $\mathbf{y}' = (y'_1, y'_2, \ldots) = h(\mathbf{x}')$. Since $\mathbf{x} \neq \mathbf{x}'$, there must an $i \in \{1, \ldots, n\}$ where $x_i \neq x'_i$. Then we have $y_i = x_i \neq x'_i = y'_i$ by the definition of h since obviously $1 \leq i \leq n$. It then follows that $h(\mathbf{x}) = \mathbf{y} = (y_1, y_2, \ldots) \neq (y'_1, y'_2, \ldots) = \mathbf{y}' = h(\mathbf{x}')$, which shows that h is injective since $\mathbf{x}$ and $\mathbf{x}'$ were arbitrary. □

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

Proof. Consider any $\mathbf{x} = (\mathbf{a}, \mathbf{b}) \in X^n \times X^\omega$ so that clearly $\mathbf{a} = (a_1, \ldots, a_n) \in X^n$ and $\mathbf{b} = (b_1, b_2, \ldots) \in X^\omega$. Then define the sequence

$$y_i = \begin{cases} a_i & 1 \leq i \leq n \\ b_{i-n} & n < i \end{cases}$$

for any $i \in \mathbb{Z}_+$, noting that when $n < i$ we have $n + 1 \leq i$ so that $1 \leq i - n$ so that b_{i-n} is defined. We then of course set $k(\mathbf{x}) = \mathbf{y} = (y_1, y_2, \ldots)$ so that clearly $k(\mathbf{x}) \in X^\omega$. Therefore k is a function from $X^n \times X^\omega$ to X^ω.

To show that k is injective consider $\mathbf{x}$ and $\mathbf{x}'$ in $X^n \times X^\omega$ where $\mathbf{x} \neq \mathbf{x}'$. Of course we have $\mathbf{x} = (\mathbf{a}, \mathbf{b})$ and $\mathbf{x}' = (\mathbf{a}', \mathbf{b}')$ where $\mathbf{a}, \mathbf{a}' \in X^n$ while $\mathbf{b}, \mathbf{b}' \in X^\omega$. It then follows that $\mathbf{a} = (a_1, \ldots, a_n)$, $\mathbf{a}' = (a'_1, \ldots, a'_n)$, $\mathbf{b} = (b_1, b_2, \ldots)$, and $\mathbf{b}' = (b'_1, b'_2, \ldots)$, where every a_i, a'_i, b_j, and b'_j are in X (for $i \in \{1, \ldots, n\}$ and $j \in \mathbb{Z}_+$). Also, let $\mathbf{y} = (y_1, y_2, \ldots) = k(\mathbf{x})$ and $\mathbf{y}' = (y'_1, y'_2, \ldots) = k(\mathbf{x}')$. Now, since $\mathbf{x} \neq \mathbf{x}'$, we have that either $\mathbf{a} \neq \mathbf{a}'$ or $\mathbf{b} \neq \mathbf{b}'$. If $\mathbf{a} \neq \mathbf{a}'$ then there is an $i \in \{1, \ldots, n\}$ where $a_i \neq a'_i$. We then have that $y_i = a_i \neq a'_i = y'_i$ by the definition of k, since obviously $1 \leq i \leq n$. If, on the other hand, $\mathbf{b} \neq \mathbf{b}'$, then there is an $i \in \mathbb{Z}_+$ such that $b_i \neq b'_i$. Then clearly $n < n + i$ since $0 < i$ so that $y_{n+i} = b_{(n+i)-n} = b_i \neq b'_i = b'_{(n+i)-n} = y'_{n+i}$, noting that clearly $n + i \in \mathbb{Z}_+$. Hence in either case there is an $i \in \mathbb{Z}_+$ such that $y_i \neq y_i'$ so that $k(\mathbf{x}) = \mathbf{y} = (y_1, y_2, \ldots) \neq (y_1', y_2', \ldots) = \mathbf{y}' = k(\mathbf{x}')$. This shows that k is injective since $\mathbf{x}$ and $\mathbf{x}'$ were arbitrary.

Now consider any $\mathbf{y} = (y_1, y_2, \ldots) \in X^\omega$ and set $a_i = y_i$ for any $i \in \{1, \ldots, n\}$ so that clearly $\mathbf{a} = (a_1, \ldots, a_n) \in X^n$. Also, for any $j \in \mathbb{Z}_+$, let $b_j = y_{n+j}$ so that clearly $\mathbf{b} = (b_1, b_2, \ldots) \in X^\omega$. Let $\mathbf{x} = (\mathbf{a}, \mathbf{b})$ so that clearly $x \in X^n \times X^\omega$. Now set $\mathbf{y}' = (y_1', y_2', \ldots) = k(\mathbf{x})$ as defined above. Consider any $i \in \mathbb{Z}_+$. If $1 \leq i \leq n$ then $y_i' = a_i = y_i$ by the definition of k. If $n < i$ then $y_i' = b_{i-n} = y_{n+(i-n)} = y_i$. Hence $y_i' = y_i$ for every $i \in \mathbb{Z}_+$ so that $k(\mathbf{x}) = \mathbf{y}' = (y_1', y_2', \ldots) = (y_1, y_2, \ldots) = \mathbf{y}$, which shows that k is surjective since $\mathbf{y}$ was arbitrary.

This completes the proof that k is bijective. □

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

Proof. Consider any $\mathbf{x} = (\mathbf{a}, \mathbf{b}) \in X^\omega \times X^\omega$ so that clearly $\mathbf{a} = (a_1, a_2, \ldots)$ and $\mathbf{b} = (b_1, b_2, \ldots)$. Now define

$$y_i = \begin{cases} a_{i/2} & i \text{ is even} \\ b_{(i+1)/2} & i \text{ is odd} \end{cases}$$

for any $i \in \mathbb{Z}_+$. Note that $i/2$ and $(i+1)/2$ are in $\mathbb{Z}_+$ if i is even or odd, respectively by Lemma 5.2.1 so that y_i is defined. Clearly we have that $y_i \in X$ for any $i \in \mathbb{Z}_+$ so that $\mathbf{y} = (y_1, y_2, \ldots) \in X^\omega$. Setting $l(\mathbf{x}) = \mathbf{y}$, we then have that l is a function from $X^\omega \times X^\omega$ to X^ω.

To show that l is injective, consider $\mathbf{x} = (\mathbf{a}, \mathbf{b})$ and $\mathbf{x}' = (\mathbf{a}', \mathbf{b}')$ in $X^\omega \times X^\omega$ where $\mathbf{x} \neq \mathbf{x}'$. Also set $\mathbf{y} = (y_1, y_2, \ldots) = l(\mathbf{x})$ and $\mathbf{y}' = (y_1', y_2', \ldots) = l(\mathbf{x}')$. Since $\mathbf{x} \neq \mathbf{x}'$, we have that either $\mathbf{a} \neq \mathbf{a}'$ or $\mathbf{b} \neq \mathbf{b}'$. If $\mathbf{a} \neq \mathbf{a}'$ then there is an $i \in \mathbb{Z}_+$ such that $a_i \neq a_i'$. Then, since clearly $2i$ is even, we have $y_{2i} = a_{(2i)/2} = a_i \neq a_i' = a'_{(2i)/2} = y'_{2i}$. On the other hand, if $\mathbf{b} \neq \mathbf{b}'$ then there is a $j \in \mathbb{Z}_+$ where $b_j \neq b_j'$. Set $k = 2j - 1$, noting that

$$\begin{aligned} 1 &\leq j \\ 2 &\leq 2j \\ 1 &\leq 2j - 1 \\ 1 &\leq k \end{aligned}$$

so that $k \in \mathbb{Z}_+$. Clearly also $(k+1)/2 = j$. Since obviously k is odd, we have $y_k = b_{(k+1)/2} = b_j \neq b_j' = b'_{(k+1)/2} = y_k'$. Hence in both cases we have that there is a $k \in \mathbb{Z}_+$ where $y_k \neq y_k'$ so that $l(\mathbf{x}) = \mathbf{y} = (y_1, y_2, \ldots) \neq (y_1', y_2', \ldots) = \mathbf{y}' = l(\mathbf{x}')$. Since $\mathbf{x}$ and $\mathbf{x}'$ were arbitrary, this shows that l is injective.

Now consider any $\mathbf{y} = (y_1, y_2, \ldots) \in X^\omega$. For any $i \in \mathbb{Z}_+$, define $a_i = y_{2i}$ and $b_i = y_{2i-1}$, noting again that $2i - 1 \in \mathbb{Z}_+$ (and clearly $2i \in \mathbb{Z}_+$). Then set $\mathbf{a} = (a_1, a_2, \ldots)$, $\mathbf{b} = (b_1, b_2, \ldots)$, and $\mathbf{x} = (\mathbf{a}, \mathbf{b})$. Now let $\mathbf{y}' = (y_1', y_2', \ldots) = l(\mathbf{x})$ and consider any $i \in \mathbb{Z}_+$. If i is even then we have by the definition of l that $y_i' = a_{i/2} = y_{2(i/2)} = y_i$. If i is odd then let $j = (i+1)/2$ so that clearly $i = 2j - 1$. Then $y_i' = b_{(i+1)/2} = b_j = y_{2j-1} = y_i$. Hence in either case we have $y_i' = y_i$ so that $l(\mathbf{x}) = \mathbf{y}' = (y_1', y_2', \ldots) = (y_1, y_2, \ldots) = \mathbf{y}$ since i was arbitrary. Since $\mathbf{y}$ was arbitrary this shows that l is surjective.

Thus we have shown that l is bijective as desired. □

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

Proof. Consider any $\mathbf{x} \in (A^\omega)^n$ so that $\mathbf{x} = (\mathbf{x}_1, \ldots \mathbf{x}_n)$ where $\mathbf{x}_i \in A^\omega$ for any $i \in \{1, \ldots, n\}$. Then let $x_{ij} = \mathbf{x}_i(j)$ for $i \in \{1, \ldots, n\}$ and $j \in \mathbb{Z}_+$ so that clearly $x_{ij} \in A$, from which it follows that each $x_{ij} \in B$ as well since $A \subset B$. Consider any $k \in \mathbb{Z}_+$. Since $n \neq 0$ (since $n \in \mathbb{Z}_+$), it follows from the Division Theorem from algebra that there are unique integers q and $0 \le r < n$ where $k = qn + r$. Suppose for a moment that $q < 0$ so that $q + 1 \le 0$. Then we have that $k = qn + r < qn + n = (q+1)n \le 0 \cdot n = 0 < k$ (since $k \in \mathbb{Z}_+$) since $r < n$ and $n > 0$ (so that $(q+1)n \le 0 \cdot n$ since $q + 1 \le 0$). This is of course a contradiction so that it must be that $q \ge 0$. Then set $i = r + 1 \ge 1$ and $j = q + 1 \ge 1$ so that $i \in \{1, \ldots, n\}$ and $j \in \mathbb{Z}_+$. Set $y_k = x_{ij}$ so that clearly $y_k \in B$ since x_{ij} is. It then follows that $\mathbf{y} = (y_1, y_2, \ldots) \in B^\omega$. Then set $m(\mathbf{x}) = \mathbf{y}$ so that clearly m is a function from $(A^\omega)^n$ to B^ω.

To show that m is injective, consider any $\mathbf{x}$ and $\mathbf{x}'$ in $(A^\omega)^n$ where $\mathbf{x} \ne \mathbf{x}'$. Then $\mathbf{x} = (\mathbf{x}_1, \ldots, \mathbf{x}_n)$ and $\mathbf{x}' = (\mathbf{x}'_1, \ldots, \mathbf{x}'_n)$ where each $\mathbf{x}_i$ and $\mathbf{x}'_i$ are in A^ω for $i \in \{1, \ldots, n\}$. As before set $x_{ij} = \mathbf{x}_i(j)$ and $x'_{ij} = \mathbf{x}'_i(j)$ for $i \in \{1, \ldots, n\}$ and $j \in \mathbb{Z}_+$, and also let $\mathbf{y} = m(\mathbf{x})$ and $\mathbf{y}' = m(\mathbf{x}')$. Now, since $\mathbf{x} \ne \mathbf{x}'$, there is an $i \in \{1, \ldots, n\}$ where $\mathbf{x}_i \ne \mathbf{x}'_i$. It then follows that there is a $j \in \mathbb{Z}_+$ such that $x_{ij} = \mathbf{x}_i(j) \ne \mathbf{x}'_i(j) = x'_{ij}$. Now let $k = (j-1)n + (i-1)$ so that it follows from the definition of m that $y_k = x_{ij}$ and $y'_k = x'_{ij}$ since the quotient q and remainder r are unique by the Division Theorem. Hence $y_k = x_{ij} \ne x'_{ij} = y'_k$ so that clearly $m(\mathbf{x}) = \mathbf{y} = (y_1, y_2, \ldots) \ne (y'_1, y'_2, \ldots) = \mathbf{y}' = m(\mathbf{x}')$. This shows that m is injective as desired since $\mathbf{x}$ and $\mathbf{x}'$ were arbitrary. □

Exercise 5.5

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

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

(b) $\{\mathbf{x} \mid x_i \ge i \text{ for all } i\}$.

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

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

Solution:

(a) Let $X = \{\mathbf{x} \in \mathbb{R}^\omega \mid x_i \text{ is an integer for all } i\}$ and $Y = \mathbb{Z}^\omega$, noting that $\mathbb{Z} \subset \mathbb{R}$. We claim that $X = Y$.

Proof. Consider any $\mathbf{x} \in X$ so that $x_i \in \mathbb{Z}$ for any $i \in \mathbb{Z}_+$. It is then immediately obvious that $\mathbf{x} \in \mathbb{Z}^\omega = Y$. Hence $X \subset Y$ since $\mathbf{x}$ was arbitrary.

Now consider any $\mathbf{x} \in Y = \mathbb{Z}^\omega$ so that $x_i \in \mathbb{Z}$ for every $i \in \mathbb{Z}_+$. Again it is obvious by the definition of X that $\mathbf{x} \in X$. Hence $Y \subset X$ since $\mathbf{x}$ was arbitrary. This shows that $X = Y$ as desired. □

(b) Let $X = \{\mathbf{x} \in \mathbb{R}^\omega \mid x_i \ge i \text{ for all } i\}$ and define $Y_i = \{x \in \mathbb{R} \mid x \ge i\}$ for $i \in \mathbb{Z}_+$, noting that obviously each $Y_i \subset \mathbb{R}$. Then let $Y = Y_1 \times Y_2 \times \cdots$. We claim that $X = Y$.

Proof. First consider $\mathbf{x} \in X$ so that $x_i \ge i$ for any $i \in \mathbb{Z}_+$. Then, for any $i \in \mathbb{Z}_+$ clearly $x_i \in Y_i$ by definition since $x_i \ge i$ (and also $x_i \in \mathbb{R}$). Hence it follows that $\mathbf{x} = (x_1, x_2, \ldots) \in Y_1 \times Y_2 \times \cdots = Y$. Since $\mathbf{x}$ was arbitrary, this shows that $X \subset Y$.

Now suppose that $\mathbf{x} \in Y$ so that $x_i \in Y_i$ for every $i \in \mathbb{Z}_+$. Consider any such $i \in \mathbb{Z}_+$ so that $x_i \in Y_i$. Then, by definition $x_i \ge i$. Since i was arbitrary, this shows that $\mathbf{x} \in X$ by definition. Hence $Y \subset X$ since $\mathbf{x}$ was arbitrary so that $X = Y$. □

(c) Define $X = \{\mathbf{x} \in \mathbb{R}^\omega \mid x_i \text{ is an integer for all } i \ge 100\}$. Also define $Y_i = \mathbb{R}$ when $i < 100$ and $Y_i = \mathbb{Z}$ when $i \ge 100$ (and $i \in \mathbb{Z}_+$ for both), noting that of course $Y_i \subset \mathbb{R}$ for either case. Let $Y = Y_1 \times Y_2 \times \cdots$, and we claim that $X = Y$.

Proof. Consider any $\mathbf{x} \in X$ so that $x_i \in \mathbb{Z}$ for all $i \geq 100$. Suppose $i \in \mathbb{Z}_+$. If $i < 100$ then clearly $x_i \in \mathbb{R} = Y_i$ since $\mathbf{x} \in \mathbb{R}^\omega$. If $i \geq 100$ then we have that $x_i \in \mathbb{Z} = Y_i$. Hence in either case $x_i \in Y_i$ so that $\mathbf{x} \in Y_1 \times Y_2 \times \cdots = Y$ since i was arbitrary. Since $\mathbf{x}$ was arbitrary, this shows that $X \subset Y$.

Now consider any $\mathbf{x} \in Y$ and any $i \in \mathbb{Z}_+$ where $i \geq 100$. Then $x_i \in Y_i = \mathbb{Z}$ so that x_i is an integer. From this it follows that $\mathbf{x} \in X$ by definition since obviously $\mathbf{x} \in \mathbb{R}^\omega$ (since $x_i \in Y_i = \mathbb{R}$ when $i < 100$). Hence $Y \subset X$ since $\mathbf{x}$ was arbitrary. This completes the proof that $X = Y$. $\square$

(d) We claim that $X = \{\mathbf{x} \in \mathbb{R}^\omega \mid x_2 = x_3\}$ *cannot* be expressed as the cartesian product of subsets of $\mathbb{R}$.

Proof. Suppose to the contrary that there are $X_i \subset \mathbb{R}$ for $i \in \mathbb{Z}_+$ where $X = X_1 \times X_2 \times \cdots$. Let $(a, a, \ldots)$ denote the sequence $(x_1, x_2, \ldots)$ where $x_i = a$ for all $i \in \mathbb{Z}_+$. We then have that $(1, 1, \ldots)$ and $(2, 2, \ldots)$ are both in X since clearly $x_2 = x_3$ in both. Hence we have that 1 and 2 are both in X_i for every $i \in \mathbb{Z}_+$ since $X = X_1 \times X_2 \times \cdots$. Now define

$$y_i = \begin{cases} 1 & i \neq 2 \\ 2 & i = 2 \end{cases}$$

for $i \in \mathbb{Z}_+$. Clearly $\mathbf{y} = (y_1, y_2, \ldots) \in X_1 \times X_2 \times \cdots$ since both 1 and 2 are in each X_i. However, it is also clear that $\mathbf{y} \notin X$ by definition since $y_2 = 2 \neq 1 = y_3$. This contradicts the fact that $X = X_1 \times X_2 \times \cdots$, which shows the desired result. $\square$

§6 Finite Sets

Exercise 6.1

(a) Make a list of all the injective maps

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

Show that none is bijective. (This constitutes a *direct* proof that a set A of cardinality three does not have cardinality four.)

(b) How many injective maps

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

are there? (You can see why one would not wish to try to prove *directly* that there is no bijective correspondence between these sets.)

Solution:

Lemma 6.1.1. *The number of injective mappings (i.e. the cardinality of the set of injective functions) from $\{1, \ldots, m\}$ to $\{1, \ldots, n\}$, where $m \leq n$, is equal to the number of m-permutations of n, which is*

$$\frac{n!}{(n-m)!} \,.$$

Proof. We fix n and show this for all $m \leq n$ by induction. First, for $m = 1$, the domain of the mappings is simply $\{1\}$ so that we need only choose a single element to which to map 1. Since there are n elements to choose from (since the range is $\{1, \ldots, n\}$) there are clearly

$$n = \frac{n!}{(n-1)!} = \frac{n!}{(n-m)!}$$

mappings, all of which are trivially injective.

Now suppose that $m < n$ and that there are $n!/(n-m)!$ injective mappings from $\{1, \ldots, m\}$ to $\{1, \ldots, n\}$. Consider any such mapping $(f_1, \ldots, f_m)$. Since this mapping is injective, each f_i is unique so that it uses m of the n available numbers in $\{1, \ldots, n\}$. Thus there are $n - m$ numbers to choose from to which to set f_{m+1} so that the mapping $(f_1, \ldots, f_{m+1})$ is still injective. Hence for each injective mapping $(f_1, \ldots, f_m)$ there are $n - m$ injective mappings from $\{1, \ldots, m+1\}$ to $\{1, \ldots, n\}$. Since there are $n!/(n-m)!$ such mappings by the induction hypothesis, the total number of mappings from $\{1, \ldots, m+1\}$ to $\{1, \ldots, n\}$ will be

$$\frac{n!}{(n-m)!}(n-m) = \frac{n!}{(n-m-1)!} = \frac{n!}{[n-(m+1)]!},$$

which completes the induction. ∎

Main Problem.

(a) Here we have $n = 4$ and $m = 3$ in Lemma 6.1.1 so that we expect $4!/(4-3)! = 4!/1! = 4! = 24$ injective mappings. Since the domain of each f is a section of the positive integers, these maps can be written simply as 3-tuples. They are enumerated below:

1. $(1, 2, 3)$	7. $(2, 1, 3)$	13. $(3, 1, 2)$	19. $(4, 1, 2)$
2. $(1, 2, 4)$	8. $(2, 1, 4)$	14. $(3, 1, 4)$	20. $(4, 1, 3)$
3. $(1, 3, 2)$	9. $(2, 3, 1)$	15. $(3, 2, 1)$	21. $(4, 2, 1)$
4. $(1, 3, 4)$	10. $(2, 3, 4)$	16. $(3, 2, 4)$	22. $(4, 2, 3)$
5. $(1, 4, 2)$	11. $(2, 4, 1)$	17. $(3, 4, 1)$	23. $(4, 3, 1)$
6. $(1, 4, 3)$	12. $(2, 4, 3)$	18. $(3, 4, 2)$	24. $(4, 3, 2)$

Note that they are all injective since no number is used more than once in each tuple. Also none are surjective since it is easily verified that there is always an element of $\{1, 2, 3, 4\}$ that is not in each tuple. Thus none are a bijection since they are not surjective.

(b) Here we have $n = 10$ and $m = 8$ in Lemma 6.1.1 so that there are $10!/(10-8)! = 10!/2! = 1814400$ injective mappings. That is nearly two million! Certainly a direct proof would be unfeasible by hand, but could be done by computer fairly easily.

Exercise 6.2

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

Solution:

Proof. Suppose that B is not finite and $B \subset A$ but that A *is* finite. Since $B \subset A$, either $B = A$ or B is a proper subset of A. In the former case we clearly have a contradiction since B would be finite since A is and $B = A$. In the latter case we have that there is a bijection from A to $\{1, \ldots, n\}$ for some $n \in \mathbb{Z}_+$ by definition since A is finite. Then, since B is a proper subset of A, it follows from Theorem 6.2 that there is a bijection from B to $\{1, \ldots, m\}$ for some $m < n$. However, then clearly B is finite by definition, which is also a contradiction since we know B is not finite. Hence in either case there is a contradiction so that A must not be finite. □

Exercise 6.3

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

Solution:

Proof. Let $Y = \{\mathbf{x} \in X^\omega \mid x_1 = 0\}$, which is clearly a proper subset of X^ω since, for example, $(1, 1, \ldots)$ is in X^ω but not in Y. We construct a bijective function f from X^ω to Y. So consider any $\mathbf{x} \in X^\omega$ and define

$$y_i = \begin{cases} 0 & i = 1 \\ x_{i-1} & i \neq 1 \end{cases}$$

for $i \in \mathbb{Z}_+$, noting that when $i \neq 1$ we have $i > 1$ so that $i - 1 \geq 1$ so that $y_i = x_{i-1}$ is defined. Now define $f(\mathbf{x}) = \mathbf{y} = (y_1, y_2, \ldots)$ so that clearly f is a function from X^ω to Y, since $y_1 = 0$ for any input $\mathbf{x}$.

To show that f is injective, consider $\mathbf{x}$ and $\mathbf{x}'$ in X^ω where $\mathbf{x} \neq \mathbf{x}'$, and let $\mathbf{y} = f(\mathbf{x})$ and $\mathbf{y}' = f(\mathbf{x}')$. Now, since $\mathbf{x} \neq \mathbf{x}'$, there is an $i \in \mathbb{Z}_+$ where $x_i \neq x_i'$. Since $i > 0$ (since $i \in \mathbb{Z}_+$) it follows that $i + 1 > 1$ so that $i + 1 \neq 1$. We then have by the definition of f that $y_{i+1} = x_{(i+1)-1} = x_i \neq x_i' = x'_{(i+1)-1} = y'_{i+1}$ so that clearly $f(\mathbf{x}) = \mathbf{y} \neq \mathbf{y}' = f(\mathbf{x}')$. Since $\mathbf{x}$ and $\mathbf{x}'$ were arbitrary, this shows that f is indeed injective.

Now consider any $\mathbf{y} \in Y$ so that $y_1 = 0$. Define $x_i = y_{i+1}$ for any $i \in \mathbb{Z}_+$ and let $\mathbf{x} = (x_1, x_2, \ldots)$. Then $\mathbf{x} \in X^\omega$ since clearly each $x_i = y_{i+1} \in X$. Now let $\mathbf{y}' = f(\mathbf{x})$ and consider any $i \in \mathbb{Z}_+$. If $i = 1$ then clearly $y_i' = y_1' = 0 = y_1 = y_i$ ($y_1' = 0$ since the range of f is Y). If $i \neq 1$ then $y_i' = x'_{i-1} = y_{(i-1)+1} = y_i$. Hence $y_i' = y_i$ in both cases so that $f(\mathbf{x}) = \mathbf{y}' = \mathbf{y}$ since i was arbitrary. This shows that f is surjective since $\mathbf{y}$ was arbitrary.

Therefore f is bijective as desired. □

Exercise 6.4

Let A be a nonempty finite simply ordered set.

(a) Show that A has a largest element. [Hint: Proceed by induction on the cardinality of A.]

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

Solution:

(a)

Proof. We show by induction that, for all $n \in \mathbb{Z}_+$, any simply ordered set with cardinality n has a largest element. This of course shows the result since, by definition, $A \neq \varnothing$ has cardinality n for some $n \in \mathbb{Z}_+$ when A is finite.

First, suppose that A is simply ordered and has cardinality 1 so that clearly $A = \{a\}$ for some element a. It is also clear that a is trivially the largest element of A since it is the only element.

Now suppose that any simply ordered set with cardinality n has a largest element. Suppose that A is simply ordered by $\prec$ and has cardinality $n+1$. Then there is a bijection f from A to $\{1, \ldots, n+1\}$, noting that obviously f^{-1} is also a bijection. Clearly A is nonempty (since the cardinality of A is $n+1 > n > 0$) so that there is an $a \in A$. Let $A' = A - \{a\}$ so that A' has cardinality n by Lemma 6.1. Note also that clearly A' is simply ordered by $\prec$ as well (technically we must restrict $\prec$ to elements of A' so that it is really ordered by $\prec \cap (A' \times A')$). It then follows that A' has a largest element b by the induction hypothesis. Since a and b must be comparable in $\prec$ by the definition of a simple order we have the following:

Case: $a = b$. This is not possible since $b \in A'$ but clearly $a \notin A - \{a\} = A'$.

Case: $a \prec b$. We claim that b is the largest element of A. To see this, consider any $x \in A$ so that either $x = a$ or $x \in A'$. In the former case clearly $x = a \preccurlyeq b$, and in the latter $x \preccurlyeq b$ since b is the largest element of A'. This shows that b is the largest element of A since x was arbitrary.

Case: $b \prec a$. We claim that a is the largest element of A. So consider any $x \in A$ so that $x = a$ or $x \in A'$. In the first case obviously $x \preccurlyeq x = a$, and in the second $x \preccurlyeq b \preccurlyeq a$ since b is the largest element of A'. This shows that a is the largest element of A since x was arbitrary.

Thus in all cases we have shown that A has a largest element, which completes the induction. □

(b)

Proof. We again show this by induction on the (finite) cardinality of the set. First, if A is a simply ordered set with cardinality 1 then clearly $A = \{a\}$ for some a, which is clearly trivially the same order type as the section $\{1\}$.

Now suppose that all simply ordered sets of cardinality n have the order type of a section of positive integers. Consider then a set A simply ordered by $\prec$ that has cardinality $n+1$. Clearly $A \neq \varnothing$ so that it has a largest element a by part (a). Then the set $A' = A - \{a\}$ has cardinality n by Lemma 6.1. Since A' is also clearly simply ordered by $\prec$ (with the appropriate restriction) it follows from the induction hypothesis that it has order type of $\{1, \ldots, m\}$ for some $m \in \mathbb{Z}_+$. Since this also implies that A' has the cardinality of m, it has to be that $m = n$ since this cardinality is unique (by Lemma 6.5). So let f' be the order-preserving bijection from A' to $\{1, \ldots, m\} = \{1, \ldots, n\}$. Now define

$$f(x) = \begin{cases} f'(x) & x \neq a \\ n+1 & x = a \end{cases}$$

for any $x \in A$. It is clear that f is a function from A to $\{1, \ldots, n+1\}$ since obviously $n+1 \in \{1, \ldots, n+1\}$ and the range of f' is $\{1, \ldots, n\} \subset \{1, \ldots, n+1\}$.

Consider next any x and x' in A where $x \prec x'$. Suppose for the moment that $x = a$. Then $x' \preccurlyeq a = x$ since a is the largest element of A. This contradicts the fact that $x \prec x'$ so that it must be that $x \neq a$. Then $f(x) = f'(x)$. If also $x' \neq a$ then clearly $f(x) = f'(x) < f'(x') = f(x')$ since f' preserves order. If $x' = a$ then $f(x') = n+1$ so that $f(x) = f'(x) \leq n < n+1 = f(x')$ since the range of f' is only $\{1, \ldots, n\}$. Hence in all cases $f(x) < f(x')$ so that f preserves order since x and x' were arbitrary. Note that this also shows that f is injective since, for any $x, x' \in A$ where $x \neq x'$, we can assume without loss of generality that $x \prec x'$ (since it must be that $x \prec x'$ or $x' \prec x$) so that $f(x) < f(x')$, and hence $f(x) \neq f(x')$.

Lastly consider any $k \in \{1, \ldots, n+1\}$. If $k = n+1$ then clearly by definition $f(a) = n+1 = k$, noting that obviously $a \in A$. On the other hand, if $k \neq n+1$ then it has to be that $k < n+1$ so $k \leq n$. Then $k \in \{1, \ldots, n\}$, which is the range of f' so that there is an $x \in A'$ where $f'(x) = k$

since f' is bijective (and therefore surjective). Since $x \in A'$ we have that $x \in A$ but $x \neq a$ so that $f(x) = f'(x) = k$. This shows that f is surjective since k was arbitrary.

Thus we have shown that f is an order-preserving bijection from A to $\{1, \ldots, n+1\}$, which completes the induction since by definition A has order type $\{1, \ldots, n+1\}$. □

Exercise 6.5

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

Solution:

We claim that in general this does not follow.

Proof. As a counterexample, let $A = \mathbb{Z}_+$ and $B = \varnothing$. Clearly A is infinite by Corollary 6.4 so that not both A and B are finite. It also follows from Exercise 5.3c that $A \times B = \varnothing$ since B is empty. Hence clearly $A \times B$ is finite. □

If we add the additional stipulation that both A and B are nonempty, then the statement becomes true.

Proof. Since $A \times B$ is finite there is a bijective function $f : A \times B \to \{1, \ldots, n\}$ for some $n \in \mathbb{Z}_+$. We then show that A is finite by first constructing an injective function g from A to $A \times B$. Since $B \neq \varnothing$, there is a $b \in B$. So, for any $x \in A$, set $g(x) = (x, b)$, which is clearly in $A \times B$ so that g is a function from A to $A \times B$. Now consider x and x' in A where $x \neq x'$. Then clearly $g(x) = (x, b) \neq (x', b) = g(x')$. This shows that g is injective since x and x' were arbitrary.

We then have that the composition $f \circ g$ is an injective function from A to $\{1, \ldots, n\}$ by Exercise 2.4b since f is injective as well (since it is a bijection). Therefore A is finite by Corollary 6.7. An analogous argument uses the fact that $A \neq \varnothing$ to show that B is also finite. □

Exercise 6.6

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

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

Solution:

(a)

Proof. We construct a bijection $f : \mathcal{P}(A) \to X^n$. So, for any $Y \in \mathcal{P}(A)$ we have that clearly $Y \subset A$. Then set

$$x_i = \begin{cases} 0 & i \notin Y \\ 1 & i \in Y \end{cases}$$

for any $i \in \{1, \ldots, n\} = A$. Now set $f(Y) = \mathbf{x} = (x_1, \ldots, x_n)$, noting that clearly $f(Y) \in X^n$ since each $x_i \in \{0, 1\} = X$. Hence f is a function from $\mathcal{P}(A)$ to X^n.

To show that f is injective consider Y and Y' in $\mathcal{P}(A)$ where $Y \neq Y'$. Also let $\mathbf{x} = f(Y)$ and $\mathbf{x}' = f(Y')$ as defined above. Since $Y \neq Y'$, we can without loss of generality assume that there is an $i \in Y$ where $i \notin Y'$. It then follows that $x_i = 1 \neq 0 = x'_i$ by the definition of f. Hence clearly $f(Y) = \mathbf{x} = (x_1, \dots, x_n) \neq (x'_1, \dots, x'_n) = \mathbf{x}' = f(Y')$, which shows that f is injective since Y and Y' were arbitrary.

Now consider any $\mathbf{x} \in X^n$ and let $Y = \{i \in A \mid x_i = 1\}$. Clearly $Y \subset A$ so that $Y \in \mathcal{P}(A)$. Let $\mathbf{x}' = f(Y)$ and consider any $i \in \{1, \dots, n\} = A$. If $i \in Y$ then $x_i = 1 = x'_i$ by the definitions of Y and f. It $i \notin Y$ then $x_i \neq 1$ so that it has to be that $x_i = 0$ since $x_i \in X = \{0, 1\}$. Also, by the definition of f, we have that $x'_i = 0 = x_i$. Thus in either case $x_i = x'_i$ so that $\mathbf{x} = \mathbf{x}' = f(Y)$ since i was arbitrary. Since $\mathbf{x}$ was arbitrary, this shows that f is surjective.

Therefore f is a bijection from A to X^n as desired. □

(b)

Proof. First, if $A = \varnothing$ then clearly $\mathcal{P}(A) = \mathcal{P}(\varnothing) = \{\varnothing\}$ is finite. So assume in what follows that $A \neq \varnothing$. Since A is finite and nonempty there is a bijection f from A to $B = \{1, \dots, n\}$ for some $n \in \mathbb{Z}_+$. Let $X = \{0, 1\}$ so that by part (a) there is a bijection g from $\mathcal{P}(B)$ to X^n. For any $Y \in \mathcal{P}(A)$ clearly the mapping $h(Y) = \left\{i \in B \mid f^{-1}(i) \in Y\right\}$ is a bijection from $\mathcal{P}(A)$ to $\mathcal{P}(B)$. It then follows that $g \circ h$ is bijection from $\mathcal{P}(A)$ to X^n. Since clearly X^n is a finite cartesian product of finite sets, it follows from Corollary 6.8 that X^n is finite so that $\mathcal{P}(A)$ must be as well since there is a bijection between them. □

Exercise 6.7

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

Solution:

Proof. As is customary, denote the set of all functions from A to B by B^A. First, if $A = \varnothing$, then the only function from A to B is the vacuous function $\varnothing$ so that $B^A = \{\varnothing\}$, which is clearly finite. So assume that $A \neq \varnothing$. Then, since A is finite, there is a bijection f from A to $\{1, \dots, n\}$ for some $n \in \mathbb{Z}_+$, noting that of course f^{-1} is then a bijection from $\{1, \dots, n\}$ to A.

We construct a bijection h from B^A to B^n. So, for any $g \in B^A$ set $h(g) = g \circ f^{-1}$, noting that clearly this is a function from $\{1, \dots, n\}$ to B. Hence h is a function from B^A to B^n.

To show that h is injective consider g and g' in B^A where $g \neq g'$. It then follows that there is an $a \in A$ where $g(a) \neq g'(a)$. Then let $k = f(a)$ so that clearly $f^{-1}(k) = a$ and $k \in \{1, \dots, n\}$. We then have that $(g \circ f^{-1})(k) = g(f^{-1}(k)) = g(a) \neq g'(a) = g'(f^{-1}(k)) = (g' \circ f^{-1})(k)$ so that it must be that $h(g) = g \circ f^{-1} \neq g' \circ f^{-1} = h(g')$. Since g and g' were arbitrary, this shows that h is injective.

Now consider any function $i \in B^n$ and let $g = i \circ f$ so that clearly g is a function from A to B since $f : A \to \{1, \dots, n\}$ and $i : \{1, \dots, n\} \to B$. Hence $g \in B^A$, and $h(g) = g \circ f^{-1} = (i \circ f) \circ f^{-1} = i \circ (f \circ f^{-1}) = i$. Since i was arbitrary, this shows that h is surjective as well.

Hence h is bijection from B^A to B^n. Now, since B^n is a finite cartesian product of finite sets (since B is finite), it is finite by Corollary 6.8. Thus it must be that B^A is also finite since there is bijection between them. □

§7 Countable and Uncountable Sets

Exercise 7.1

Show that $\mathbb{Q}$ is countably infinite.

Solution:

Lemma 7.1.1. *The set $\mathbb{Z} \times \mathbb{Z}$ is countably infinite.*

Proof. First, by Example 7.1, the set of integers $\mathbb{Z}$ is countably infinite so that there is a bijection f from $\mathbb{Z}$ to $\mathbb{Z}_+$. We construct a bijection g from $\mathbb{Z} \times \mathbb{Z}$ to $\mathbb{Z}_+ \times \mathbb{Z}_+$. For any $(a, b) \in \mathbb{Z} \times \mathbb{Z}$ define $g(a, b) = (f(a), f(b))$, noting that clearly $g(a, b) \in \mathbb{Z}_+ \times \mathbb{Z}_+$ since $\mathbb{Z}_+$ is the range of f. Hence g is a function from $\mathbb{Z} \times \mathbb{Z}$ to $\mathbb{Z}_+ \times \mathbb{Z}_+$.

It is easy to show that g is bijective. First, consider any (a, b) and (a', b') in $\mathbb{Z}\times\mathbb{Z}$ where $(a, b) \neq (a', b')$ so that $a \neq a'$ or $b \neq b'$. If $a \neq a'$ then $f(a) \neq f(a')$ since f is bijective (and therefore injective). Thus we have that $g(a, b) = (f(a), f(b)) \neq (f(a'), f(b')) = g(a', b')$. A similar argument shows the same result when $b \neq b'$. Since (a, b) and (a', b') were arbitrary, this shows that g is injective.

Now consider any $(c, d) \in \mathbb{Z}_+ \times \mathbb{Z}_+$ so that $c, d \in \mathbb{Z}_+$. Since f is surjective (since it is a bijection) there are $a, b \in \mathbb{Z}$ where $f(a) = c$ and $f(b) = d$. We then clearly have that $g(a, b) = (f(a), f(b)) = (c, d)$ so that g is surjective (c, d) was arbitrary.

Therefore g is a bijection. Now, we know from Corollary 7.4 that $\mathbb{Z}_+ \times \mathbb{Z}_+$ is countably infinite so that there must be a bijection h from $\mathbb{Z}_+ \times \mathbb{Z}_+$ to $\mathbb{Z}_+$. It then follows that $h \circ g$ is bijection from $\mathbb{Z} \times \mathbb{Z}$ to $\mathbb{Z}_+$, which shows the desired result by definition. □

Main Problem.

Proof. First we define a straightforward function f from $\mathbb{Z}\times\mathbb{Z}$ to $\mathbb{Q}$. First consider any $(m, n) \in \mathbb{Z}\times\mathbb{Z}$. If $n \neq 0$ then let $q = m/n$. If $n = 0$ then set $q = 0$. Setting $f(m, n) = q$ we clearly have that f is a function from $\mathbb{Z} \times \mathbb{Z}$ to $\mathbb{Q}$. Now consider any rational q so that by definition there are integers m and n where $q = m/n$. It then of course follows that $f(m, n) = m/n = q$, which shows that f is surjective since q was arbitrary.

Now, from Lemma 7.1.1 we know that $\mathbb{Z} \times \mathbb{Z}$ is countably infinite so that there is a bijection g from $\mathbb{Z} \times \mathbb{Z}$ to $\mathbb{Z}_+$. Hence g^{-1} is a bijection from $\mathbb{Z}_+$ to $\mathbb{Z} \times \mathbb{Z}$. It then follows that the function $f \circ g^{-1}$ is a surjective function from $\mathbb{Z}_+$ to $\mathbb{Q}$. From this it follows from Theorem 7.1 that $\mathbb{Q}$ is countable. Since $\mathbb{Z}_+$ is a subset of $\mathbb{Q}$, it has to be that $\mathbb{Q}$ is infinite, and hence must be countably infinite. □

Exercise 7.2

Show that the maps f and g of Examples 1 and 2 are bijections.

Solution:

It is claimed in Example 7.1 that the function

$$f(n) = \begin{cases} 2n & n > 0 \\ -2n + 1 & n \leq 0 \end{cases}$$

is a bijection from $\mathbb{Z}$ to $\mathbb{Z}_+$.

Proof. To show that f is injective, consider $n, m \in \mathbb{Z}$ where $n \neq m$.

Case: $n > 0$. Then $f(n) = 2n$, which is clearly even. If $m > 0$, then clearly $f(n) = 2n \neq 2m = f(m)$ since $n \neq m$. If $m \leq 0$ then $f(m) = -2m + 1 = 2(-m) + 1$ is clearly odd so that it must be that $f(n) \neq f(m)$.

Case: $n \leq 0$. Then $f(n) = -2n + 1 = 2(-n) + 1$, which is clearly odd. If $m > 0$ then $f(m) = 2m$ is even so that it has to be that $f(n) \neq f(m)$. If $m \leq 0$ then $f(m) = -2m + 1 \neq -2n + 1 = f(n)$ since $n \neq m$.

Thus in every case $f(n) \neq f(m)$, which shows that f is injective since n and m were arbitrary.

To show that f is surjective, consider any $k \in \mathbb{Z}_+$. If k is even then $k = 2n$ for some $n \in \mathbb{Z}_+$. Hence $n > 0$ (since $k > 0$ and $n = k/2$) so that $f(n) = 2n = k$, noting that $n \in \mathbb{Z}$ since $\mathbb{Z}_+ \subset \mathbb{Z}$. If k is odd then $k = 2m - 1$ for some $m \in \mathbb{Z}_+$. So let $n = 1 - m$ so that clearly n is an integer and

$$\begin{aligned} m &\geq 1 && \text{(since } m \in \mathbb{Z}_+\text{)} \\ -m &\leq -1 \\ 1 - m &\leq 0 \\ n &\leq 0\,. \end{aligned}$$

Thus $f(n) = -2n + 1 = -2(1 - m) + 1 = -2 + 2m + 1 = 2m - 1 = k$. This shows that f is surjective since k was arbitrary. Therefore we have shown that f is a bijection as desired. □

Regarding Example 7.2, the following set is defined:

$$A = \{(x, y) \in \mathbb{Z}_+ \times \mathbb{Z}_+ \mid y \leq x\}\ .$$

Then the function f is defined from $\mathbb{Z}_+ \times \mathbb{Z}_+$ to A by

$$f(x, y) = (x + y - 1, y)$$

for $(x, y) \in \mathbb{Z}_+ \times \mathbb{Z}_+$. It is claimed that f is a bijection.

Proof. First, it is not even clear that the range of f is constrained to A, so let us show this. Consider any $(x, y) \in \mathbb{Z}_+ \times \mathbb{Z}_+$ so that $f(x, y) = (x + y - 1, y)$. Since $x \geq 1$ and $y \geq 1$, we have that $x + y \geq 1 + 1 = 2 > 1$ so that $x + y - 1 > 0$ and hence $x + y - 1 \in \mathbb{Z}_+$. Thus clearly $f(x, y) = (x + y - 1, y) \in \mathbb{Z}_+ \times \mathbb{Z}_+$. We also have

$$\begin{aligned} 1 &\leq x \\ 0 &\leq x - 1 \\ y &\leq x + y - 1\,. \end{aligned}$$

Therefore it is clear that $f(x, y) = (x + y - 1, y) \in A$ by definition.

To show that f is injective consider (x, y) and (x', y') in $\mathbb{Z}_+ \times \mathbb{Z}_+$ where $f(x, y) = (x + y - 1, y) = (x' + y' - 1, y') = f(x', y')$. Thus $x + y - 1 = x' + y' - 1$ and $y = y'$. Therefore $x + y - 1 = x' + y' - 1 = x' + y - 1$, from which it obviously follows that $x = x'$ as well. Then $(x, y) = (x', y')$, which shows that f is injective since (x, y) and (x', y') were arbitrary.

Now consider any $(z, y) \in A$ so that $(z, y) \in \mathbb{Z}_+ \times \mathbb{Z}_+$ and $y \leq z$. Let $x = z - y + 1$ so that clearly $z = x + y - 1$. We also have

$$\begin{aligned} y &\leq z = x + y - 1 \\ 0 &\leq x - 1 \\ 1 &\leq x \end{aligned}$$

so that $(x, y) \in \mathbb{Z}_+ \times \mathbb{Z}_+$. Since also we have $f(x, y) = (x + y - 1, y) = (z, y)$, f is surjective since (z, y) was arbitrary. This completes the proof that f is a bijection. □

The function g is then defined from A to $\mathbb{Z}_+$ by

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

for $(x, y) \in A$. This is also claimed to be a bijection.

Proof. First we show that the range of g is indeed $\mathbb{Z}_+$ since this is not obvious. Consider any $(x, y) \in A$ so that $(x, y) \in \mathbb{Z}_+ \times \mathbb{Z}_+$ and $y \leq x$. First, if x is even then $x = 2n$ for some $n \in \mathbb{Z}$. Then $g(x, y) = (x-1)x/2 + y = (2n-1)(2n)/2 + y = (2n-1)n + y$, which is clearly an integer. If x is odd then $x = 2n+1$ for some integer n so that

$$g(x,y) = (x-1)x/2 + y = (2n+1-1)(2n+1)/2 + y = (2n)(2n+1)/2 + y = n(2n+1) + y\,,$$

which is also clearly an integer. We also have that $-y < 0$ since $y > 0$ so that

$$\begin{aligned}
x &\geq 1 \\
x - 1 &\geq 0 \\
\frac{1}{2}(x-1) &\geq 0 && \text{(since } 1/2 > 0) \\
\frac{1}{2}(x-1)x &\geq 0 > -y && \text{(since } x > 0) \\
\frac{1}{2}(x-1)x + y &> 0 \\
g(x,y) &> 0\,.
\end{aligned}$$

Since we have shown that $g(x, y) \in \mathbb{Z}$ as well, it follows that $g(x, y) \in \mathbb{Z}_+$.

Consider any $(x, y) \in A$ so that $(x, y) \in \mathbb{Z}_+ \times \mathbb{Z}_+$ and $y \leq x$. Then clearly

$$\begin{aligned}
g(x,y) &= \frac{1}{2}(x-1)x + y \leq \frac{1}{2}(x-1)x + x \\
&< \frac{1}{2}(x-1)x + x + 1 = \frac{1}{2}(x^2 - x + 2x) + 1 \\
&= \frac{1}{2}(x^2 + x) + 1 = \frac{1}{2}x(x+1) + 1 \\
&= \frac{1}{2}(x+1-1)(x+1) + 1 \\
&= g(x+1, 1)\,.
\end{aligned}$$

A simple inductive argument shows that $g(x, y) < g(x+n, 1)$ for any $n \in \mathbb{Z}_+$. This was just shown for $n = 1$. Then, assuming it true for n, we have that $g(x, y) < g(x+n, 1) < g((x+n)+1, 1) = g(x+(n+1), 1)$, which completes the induction.

So consider any (x, y) and (x', y') in A so that (x, y) and (x', y') are in $\mathbb{Z}_+ \times \mathbb{Z}_+$, $y \leq x$, and $y' \leq x'$. Also suppose that $(x, y) \neq (x', y')$ so that either $x \neq x'$ or $y \neq y$. If $x = x'$ then it has to be that $y \neq y'$ so that clearly

$$\begin{aligned}
y &\neq y' \\
\frac{1}{2}(x-1)x + y &\neq \frac{1}{2}(x-1)x + y' \\
\frac{1}{2}(x-1)x + y &\neq \frac{1}{2}(x'-1)x' + y' \\
g(x,y) &\neq g(x',y')\,.
\end{aligned}$$

If $x \neq x'$ then we can assume that $x < x'$. Then let $n = x' - x$ so that clearly $n > 0$ and $x' = x + n$. By what was just shown, we have

$$g(x,y) < g(x+n, 1) = g(x', 1) = \frac{1}{2}(x'-1)x' + 1 \leq \frac{1}{2}(x'-1)x' + y' = g(x', y')$$

since $1 \leq y'$. Thus $g(x, y) \neq g(x', y')$. Since this is true in both cases, this shows that g is injective since (x, y) and (x', y') were arbitrary.

To show that g is also surjective, consider any $z \in \mathbb{Z}_+$. Define the set $B = \{x \in \mathbb{Z}_+ \mid g(x, 1) \leq z\}$. First, we have that $g(1, 1) = 1 \leq z$ since $z \in \mathbb{Z}_+$ so that $1 \in B$ and therefore $B \neq \varnothing$. If $z = 1$ then clearly $z = 1 \leq 1 = g(1, 1) = g(z, 1)$. If $z \neq 1$ then we have

$$\begin{aligned}
2 &\leq z \\
1 &\leq \frac{1}{2}z \\
z - 1 &\leq \frac{1}{2}(z-1)z \\
z &\leq \frac{1}{2}(z-1)z + 1 \\
z &\leq g(z, 1)
\end{aligned}$$

Now consider any $x, y \in \mathbb{Z}_+$ where $x < y$. It then follows from what was shown above that $g(x, 1) \leq g(x, y) < g(x + 1, 1)$. From this we clearly have that the function $g(x, 1)$ is monotonically increasing in x, i.e. for $x, y \in \mathbb{Z}_+$, $x < y$ implies that $g(x, 1) < g(y, 1)$. By the contrapositive of this, $g(x, 1) \geq g(y, 1)$ implies that $x \geq y$. With this in mind, consider any $x \in B$ so $g(x, 1) \leq z \leq g(z, 1)$. Then this implies that $x \leq z$, which shows that z is an upper bound of B since x was arbitrary.

We have thus shown that B is a nonempty set of integers that is bounded above. It then follows from Exercise 4.9a that B has a largest element x. Now let $y = z - g(x, 1) + 1$, noting that, since $x \in B$,

$$\begin{aligned}
g(x, 1) &\leq z \\
0 &\leq z - g(x, 1) \\
1 &\leq z - g(x, 1) + 1 \\
1 &\leq y
\end{aligned}$$

and hence $y \in \mathbb{Z}_+$ so that $(x, y) \in \mathbb{Z}_+ \times \mathbb{Z}_+$. We also must have that $z < g(x + 1, 1)$ since otherwise we would have that $x + 1 \in B$, which would violate the definition of x as being the largest element of B. Thus we have

$$\begin{aligned}
z &\leq g(x + 1, 1) - 1 \\
z &\leq \frac{1}{2}(x + 1 - 1)(x + 1) + 1 - 1 \\
z &\leq \frac{1}{2}(x + 1)x \\
z &\leq x + \frac{1}{2}(x - 1)x \\
z &\leq x + \frac{1}{2}(x - 1)x + 1 - 1 \\
z &\leq x + g(x, 1) - 1 \\
z - g(x, 1) + 1 &\leq x \\
y &\leq x
\end{aligned}$$

so that $(x, y) \in A$.

Lastly, since $y = z - g(x, 1) + 1$, we clearly have

$$z = y + g(x, 1) - 1 = y + \frac{1}{2}(x - 1)x + 1 - 1 = \frac{1}{2}(x - 1)x + y = g(x, y)\,.$$

This shows that g is surjective since z was arbitrary, thereby completing the long and arduous proof that g is a bijection. $\square$

Exercise 7.3

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:

Proof. Similar to Exercise 6.6a, we construct such a bijection f from $\mathcal{P}(\mathbb{Z}_+)$ to X^ω. For any $A \in \mathcal{P}(\mathbb{Z}_+)$ we have that $A \subset \mathbb{Z}_+$. Then, for $i \in \mathbb{Z}_+$, set

$$x_i = \begin{cases} 1 & i \in A \\ 0 & i \notin A \end{cases}$$

and set $f(A) = (x_1, x_2, \ldots)$ so that clearly $f(A) \in X^\omega$.

To show that f is injective consider A and A' in $\mathcal{P}(\mathbb{Z}_+)$ where $A \neq A'$. Without loss of generality, we can assume that there is an $i \in A$ where $i \notin A'$, noting that of course $i \in \mathbb{Z}_+$ since $A \subset \mathbb{Z}_+$. Let $\mathbf{x} = (x_1, x_2, \ldots) = f(A)$ and $\mathbf{x}' = (x_1', x_2', \ldots) = f(A')$. Then $x_i = 1 \neq 0 = x_i'$ by the definition of f since $i \in A$ but $i \notin A'$. Thus clearly $f(A) = \mathbf{x} \neq \mathbf{x}' = f(A')$, which shows that f is injective since A and A' were arbitrary.

Now consider any $\mathbf{x} = (x_1, x_2, \ldots) \in X^\omega$ and define the set $A = \{i \in \mathbb{Z}_+ \mid x_i = 1\}$ so that clearly $A \subset \mathbb{Z}_+$ and hence $A \in \mathcal{P}(\mathbb{Z}_+)$. Let $\mathbf{x}' = (x_1', x_2', \ldots) = f(A)$ and consider $i \in \mathbb{Z}_+$. If $i \in A$ then $x_i' = 1 = x_i$ by the definitions of A and f. If $i \notin A$ then $x_i \neq 1$ since otherwise $i \in A$ by definition. Since $x_i \in X = \{0, 1\}$ it clearly must be that $x_i = 0$. We then also have that $x_i' = 0$ by the definition of f, and thus $x_i = 0 = x_i'$. Since $x_i = x_i'$ in both cases and i was arbitrary, it follows that $\mathbf{x} = \mathbf{x}' = f(A)$. This proves that f is surjective since $\mathbf{x}$ was arbitrary.

Hence it has been shown that f is a bijection as desired. $\square$

Exercise 7.4

(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. (It is a somewhat surprising fact that only two transcendental numbers are familiar to us: e and π. Even proving these two numbers transcendental is highly nontrivial.)

Solution:

(a)

Proof. First consider arbitrary degree $n \in \mathbb{Z}_+$. Then for each $\mathbf{q} = (q_1, \ldots, q_n) \in \mathbb{Q}^n$, there is a corresponding polynomial equation in x:

$$x^n + \sum_{i=1}^{n} q_i x^{i-1} = x^n + q_n x^{n-1} + \cdots + q_2 x + q_1 = 0\,,$$

which is assumed to have a finite number of solutions. So let $X_{\mathbf{q}}$ be the finite set real numbers that are solutions. (We note that the polynomial corresponding to the vector $\mathbf{q} = (0, \ldots, 0) \in \mathbb{Q}^n$ becomes $0 = 0$ so that any real number x satisfies it. Similarly the polynomial corresponding to $\mathbf{q} = (q_1, 0, \ldots, 0) \in \mathbb{Q}^n$ for nonzero q_1 corresponds to the equation $q_1 = 0$, which has no solutions. Of course $X_{\mathbf{q}} = \varnothing$ is still finite in this case. For the infinite-solution case we could simply remove the zero vector from $\mathbb{Q}^n$ without changing the argument in any substantial way. This is also taken care of if we really do assume that *any* polynomial has a finite number of solutions as we are evidently doing here.)

Now, we clearly have that $\mathbb{Q}^n$ is countable by Theorem 7.6 since it is a finite product of countable sets (since it was shown in Exercise 7.1 that $\mathbb{Q}$ is countable). Thus the set $A_n = \bigcup_{\mathbf{q} \in \mathbb{Q}^n} X_{\mathbf{q}}$ is countable by Theorem 7.5 since it is a countable union of finite (and therefore countable) sets. Of course, this is the set of all algebraic numbers from polynomials of degree n. Then $A = \bigcup_{n \in \mathbb{Z}_+} A_n$ is the set of all algebraic numbers, which is also then countable by Theorem 7.5 since each A_n was shown to be countable. □

(b)

Proof. As in part (a), let $A \subset \mathbb{R}$ be the set of algebraic numbers so that clearly, by definition, $T = \mathbb{R} - A$ is the set of transcendental numbers. Note that clearly $\mathbb{R} = A \cup T$ so that, if T were countable, then $\mathbb{R}$ would be too since it is a finite union of countable sets. This of course contradicts the (hitherto unproven) fact that $\mathbb{R}$ is uncountable so that it must be that T is also uncountable as desired. □

Exercise 7.5

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 $f : \{1, \ldots, n\} \to \mathbb{Z}_+$.
(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:

(a) The set A of all functions $f : \{0, 1\} \to \mathbb{Z}_+$.

We claim that A is countable.

Proof. For any $f \in A$, clearly the mapping $g(f) = (f(0), f(1))$ is a bijection from A to $\mathbb{Z}_+{}^2$. Since $\mathbb{Z}_+{}^2$ is a finite cartesian product of countable sets, it follows that it is also countable by Theorem 7.6. Hence there is a bijection $h : \mathbb{Z}_+{}^2 \to \mathbb{Z}_+$. It is then obvious that $h \circ g$ is a bijection from A to $\mathbb{Z}_+$ so that A is countable. □

(b) The set B_n of all functions $f : \{1, \ldots, n\} \to \mathbb{Z}_+$.

We claim that B_n (for some $n \in \mathbb{Z}_+$) is also countable.

Proof. By the definition of $\mathbb{Z}_+{}^n$, $B_n = \mathbb{Z}_+{}^n$, which is clearly a finite cartesian product of countable sets. Thus B_n is countable by Theorem 7.6. □

(c) The set $C = \bigcup_{n \in \mathbb{Z}_+} B_n$.

We claim that C is countable.

Proof. Since n was arbitrary in part (b), we showed that B_n is countable for any $n \in \mathbb{Z}_+$. Thus $C = \bigcup_{n \in \mathbb{Z}_+} B_n$ is a countable union of countable sets, which is itself also countable by Theorem 7.5 as desired. □

(d) The set D of all functions $f : \mathbb{Z}_+ \to \mathbb{Z}_+$.

Clearly $D = \mathbb{Z}_+{}^\omega$, which we claim is uncountable.

Proof. We proceed to show, as in Theorem 7.7, that any function $g : \mathbb{Z}_+ \to D$ is not surjective. So denote

$$g(n) = \mathbf{x}_n = (x_{n1}, x_{n2}, \ldots),$$

where of course each $x_{nm} \in \mathbb{Z}_+$ since $\mathbf{x}_n \in D$ and so is a function from $\mathbb{Z}_+$ to $\mathbb{Z}_+$. Now set

$$y_n = \begin{cases} 0 & x_{nn} \neq 0 \\ 1 & x_{nn} = 0 \end{cases}$$

so that clearly $\mathbf{y} = (y_1, y_2, \ldots)$ is an element of D. Now consider any $n \in \mathbb{Z}_+$. If $x_{nn} = 0$ then $y_n = 1 \neq 0 = x_{nn}$, and if $x_{nn} \neq 0$ then $y_n = 0 \neq x_{nn}$. Thus clearly $g(n) = \mathbf{x}_n \neq \mathbf{y}$ since the nth element of each differs. This shows that g cannot be surjective since $\mathbf{y} \in D$ and n was arbitrary. It then follows from Theorem 7.1 that D is not countable. □

(e) The set E of all functions $f : \mathbb{Z}_+ \to \{0, 1\}$.

This is exactly the set X^ω in Theorem 7.7, wherein it was shown to be uncountable.

(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 $n \geq N$.]

We claim that F is countable.

Proof. For brevity define $X = \{0, 1\}$. First let F_N be the set of all eventually zero functions $f : \mathbb{Z}_+ \to X$ that are zero for $n \geq N$, where of course $N \in \mathbb{Z}_+$. Then clearly $F = \bigcup_{N \in \mathbb{Z}_+} F_N$.

We show that each F_N is countable. So consider any $N \in \mathbb{Z}_+$. If $N = 1$ then clearly $f : \mathbb{Z}_+ \to X$ defined by $f(n) = 0$ for $n \in \mathbb{Z}_+$ (which could be denoted $(0, 0, \ldots)$) is the only element of $F_N = F_1$ so that f_N is clearly finite and therefore countable. If $N > 1$ then for $\mathbf{x} = (x_1, \ldots, x_{N-1}) \in X^{N-1}$ define

$$y_n = \begin{cases} x_n & n < N \\ 0 & n \geq N \end{cases}$$

for $n \in \mathbb{Z}_+$. It then trivial to show that g defined by $g(\mathbf{x}) = \mathbf{y} = (y_1, y_2, \ldots)$ is a bijection from X^{N-1} to F_N. Now, since $X = \{0, 1\}$ is finite, X^{N-1} is finite by Corollary 6.8. Since this is in bijective correspondence with F_N, it follows that it must also be finite and therefore countable.

Thus $F = \bigcup_{N \in \mathbb{Z}_+} F_N$ is a countable union of countable sets, and so is countable by Theorem 7.5 as desired □

(g) The set G of all functions $f : \mathbb{Z}_+ \to \mathbb{Z}_+$ that are eventually 1.

Since G is clearly a subset of H in part (h) below, it is countable by Corollary 7.3 since H is.

(h) The set H of all functions $f : \mathbb{Z}_+ \to \mathbb{Z}_+$ that are eventually constant.

We claim that H is countable.

Proof. For $N \in \mathbb{Z}_+$, let H_N be the set of functions $f : \mathbb{Z}_+ \to \mathbb{Z}_+$ such that $f(n)$ is constant for $n \geq N$. Thus clearly $H = \bigcup_{N \in \mathbb{Z}_+} H_N$.

We show that each H_N is countable. So consider $N \in \mathbb{Z}_+$. For any $\mathbf{x} = (x_1, \ldots, x_N) \in {\mathbb{Z}_+}^N$ define

$$y_n = \begin{cases} x_n & n < N \\ x_N & n \geq N \end{cases}$$

for $n \in \mathbb{Z}_+$, and set $g(\mathbf{x}) = \mathbf{y} = (y_1, y_2, \ldots)$. It is then a simple matter to show that g is a bijection from ${\mathbb{Z}_+}^N$ to H_N. Then, since ${\mathbb{Z}_+}^N$ is a finite product of countable sets, it is countable by Theorem 7.6. Hence H_N must also be countable since there is a bijective correspondence between them.

Thus $H = \bigcup_{N \in \mathbb{Z}_+} H_N$ is the countable union of countable sets so that it must also be countable by Theorem 7.5. □

(i) The set I of all two-element subsets of $\mathbb{Z}_+$.

In part (j) below it is shown that the set J of all finite subsets of $\mathbb{Z}_+$ is countable. Since clearly $I \subset J$, it follows that I is also countable by Corollary 7.3.

(j) The set J of all finite subsets of $\mathbb{Z}_+$.

We claim that J is countable.

Proof. First, let J_n denote the set of n-element subsets of $\mathbb{Z}_+$ (for $n \in pints$), and let $J_0 = \{\varnothing\}$ since $\varnothing$ is the only "zero-element" subset of $\mathbb{Z}_+$. Clearly then $J = \bigcup_{n \in \mathbb{Z}_+ \cup \{0\}} J_n$. Obviously J_0 is finite and therefore countable. Next, we show that J_n is countable for any $n \in \mathbb{Z}_+$.

So consider any such $n \in \mathbb{Z}_+$. Clearly ${\mathbb{Z}_+}^n$ is countable by Theorem 7.6 since it is a finite product of countable sets. Hence there is a bijection $f : {\mathbb{Z}_+}^n \to \mathbb{Z}_+$. We now construct an injective function $g : J_n \to {\mathbb{Z}_+}^n$. For any $X \in J_n$, we can choose a bijection $h : X \to \{1, \ldots, n\}$ since it has n elements. Since $X \subset \mathbb{Z}_+$, clearly $h^{-1} \in {\mathbb{Z}_+}^n$, so set $g(X) = h^{-1}$. To show that g is injective, consider X and X' in J_n where $X \neq X'$. Without loss of generality we can assume that there is an $x \in X$ where $x \notin X'$. Let h and h' be the chosen bijections from X and X', respectively, to $\{1, \ldots, n\}$ so that by definition $g(X) = h^{-1}$ and $g(X') = h'^{-1}$. Now let $k = h(x)$ so that $h^{-1}(k) = x$. It has to be that $h'^{-1}(k) \neq x$ since otherwise x would be in X'. Hence $h^{-1}(k) = x \neq h'^{-1}(k)$, which shows that $g(X) = h^{-1} \neq h'^{-1} = g(X')$. Thus g is injective since X and X' were arbitrary. It then follows that $f \circ g$ is an injective function from J_n to $\mathbb{Z}_+$ so that J_n must be countable by Theorem 7.1.

Since n was arbitrary, this shows that J_n is countable for any $n \in \mathbb{Z}_+$. From this it follows from Theorem 7.5 that $J = \bigcup_{n \in \mathbb{Z}_+ \cup \{0\}} J_n$ is also countable since it is clearly a countable union of countable sets. □

Exercise 7.6

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. [Hint: Define $A_1 = A$, $B_1 = B$, and for $n > 1$, $A_n = f(A_{n-1})$ and $B_n = f(B_{n-1})$. (Recursive definition again!) Note that $A_1 \supset B_1 \supset A_2 \supset B_2 \supset A_3 \supset \cdots$. Define a bijection $h : A \to B$ by the rule

$$h(x) = \begin{cases} f(x) & \text{if } x \in A_n - B_n \text{ for some } n, \\ x & \text{otherwise.]} \end{cases}$$

(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:

(a)

Proof. Following the hint, we define two sequences of sets recursively:

$$A_1 = A \qquad\qquad B_1 = B$$

and

$$A_n = f(A_{n-1}) \qquad\qquad B_n = f(B_{n-1})$$

for integer $n > 1$. Now define a function from A to B by

$$h(x) = \begin{cases} f(x) & x \in A_n - B_n \text{ for some } n \in \mathbb{Z}_+ \\ x & \text{otherwise} \end{cases}$$

for any $x \in A$.

First we show that B really is the range of h as this is not readily apparent. So consider any $x \in A$. Clearly if $x \in A_n - B_n$ for some $n \in \mathbb{Z}_+$ then $h(x) = f(x) \in B$ since B is the range of f. On the other hand, if this is not the case then $x \notin A_n - B_n$ for any $n \in \mathbb{Z}_+$, and $h(x) = x$. In particular, $x \notin A_1 - B_1 = A - B$ so that it has to be that $h(x) = x \in B$, for otherwise it would be that $x \in A - B$ since $x \in A$. Hence, in either case, $h(x) \in B$ so that h is indeed a function from A to B.

To show that h is injective, consider any $x, x' \in A$ where $x \neq x'$.

1. Case: $x \in A_n - B_n$ for some $n \in \mathbb{Z}_+$. Then by definition $h(x) = f(x)$.
   (a) Case: $x' \in A_m - B_m$ for some $m \in \mathbb{Z}_+$. Then we clearly have $h(x) = f(x) \neq f(x') = h(x')$ since f is injective and $x \neq x'$.
   (b) Case: $x' \notin A_m - B_m$ for all $m \in \mathbb{Z}_+$. Then $h(x') = x'$. Since $x \in A_n$, we have that $f(x) \in f(A_n) = A_{n+1}$. If it were the case that $f(x) \in B_{n+1} = f(B_n)$, then there would be a $y \in B_n$ such that $f(y) = f(x)$. Of course, since f is injective, it would have to be that $x = y \in B_n$, which we know is not the case since $x \in A_n - B_n$. Hence it has to be that $f(x) \notin B_{n+1}$ so that $f(x) \in A_{n+1} - B_{n+1}$. From this it is clearly that it cannot be that $x' = f(x)$ so that $h(x') = x' \neq f(x) = h(x)$.

2. Case: $x \notin A_n - B_n$ for all $n \in \mathbb{Z}_+$. Then by definition $h(x) = x$.
   (a) Case: $x' \in A_m - B_m$ for some $m \in \mathbb{Z}_+$. This is the same as case 1b above with the roles of x and x' reversed.
   (b) Case: $x' \notin A_m - B_m$ for all $m \in \mathbb{Z}_+$. Then clearly $h(x) = x \neq x' = h(x')$.

Thus in all cases $h(x) \neq h(x')$, which shows that h is injective since x and x' were arbitrary.

To show that h is also surjective, consider any $y \in B$, noting that also $y \in A$ since $B \subset A$.

Case: $y \in A_n - B_n$ for some $n \in \mathbb{Z}_+$. It cannot be that $n = 1$ since then $y \in A_1 - B_1 = A - B$, and we know that $y \in B$. Hence $n > 1$ so that $n - 1 \in \mathbb{Z}_+$. Since $y \in A_n = f(A_{n-1})$, there is an $x \in A_{n-1}$ where $f(x) = y$. Suppose for a moment that $x \in B_{n-1}$ so that $y = f(x) \in f(B_{n-1}) = B_n$, which we know not to be the case. Thus it must be that $x \notin B_{n-1}$ so that $x \in A_{n-1} - B_{n-1}$ and so by definition $h(x) = f(x) = y$.

Case: $y \notin A_n - B_n$ for all $n \in \mathbb{Z}_+$. Then clearly $h(y) = y$ by definition.

This shows that h is surjective since y was arbitrary.

Therefore it has been shown that h is a bijection from A to B, which shows that they have the same cardinality by definition. □

(b)

Proof. Clearly f is a bijection from A to $f(A)$ since f is injective. Also, clearly the function $g \circ f$ is an injective function from C into $f(A)$ since both f and g are injective. Noting that obviously $f(A) \subset C$, it then follows from part (a) that C and $f(A)$ have the same cardinality so that there is a bijection $h : f(A) \to C$. We then have that $h \circ f$ is a bijection from A to C so that they have the same cardinality by definition. □

Exercise 7.7

Show that the sets D and E of Exercise 7.5 have the same cardinality.

Solution:

Throughout what follows let A^B denote the set of all functions from set A to set B.

Lemma 7.7.1. *If there is an injection from A_1 to A_2 with $A_2 \neq \varnothing$, and an injection from B_1 to B_2, then there is also an injection from $A_1^{B_1}$ to $A_2^{B_2}$.*

Proof. Since $A_2 \neq \varnothing$, there is an $a_2 \in A_2$. Since we know they exist, let $f_A : A_1 \to A_2$ and $f_B : B_1 \to B_2$ be injections. We construct an injection $F : A_1^{B_1} \to A_2^{B_2}$. So, for any $g \in A_1^{B_1}$, define $F(g) = h$, where $h \in A_2^{B_2}$ is defined by

$$h(b) = \begin{cases} (f_A \circ g \circ f_B^{-1})(b) & b \in f_B(B_1) \\ a_2 & b \notin f_B(B_1) \end{cases}$$

for $b \in B_2$, noting that f_B^{-1} is a function with domain $f_B(B_1)$ since it is injective.

To show that F is injective, consider $g_1, g_2 \in A_1^{B_1}$ where $g_1 \neq g_2$. Then there is a $b_1 \in B_1$ where $g_1(b_1) \neq g_2(b_1)$. Let $b_2 = f_B(b_1)$ so that clearly $b_2 \in f_B(B_1)$ and $b_1 = f_B^{-1}(b_2)$. Then clearly

$$F(g_1)(b_2) = (f_A \circ g_1 \circ f_B^{-1})(b_2) = f_A(g_1(f_B^{-1}(b_2))) = f_A(g_1(b_1))$$

$$\neq f_A(g_2(b_1)) = f_A(g_2(f_B^{-1}(b_2))) = (f_A \circ g_2 \circ f_B^{-1})(b_2)$$
$$= F(g_2)(b_2)$$

since $g_1(b_1) \neq g_2(b_1)$ and f_A is injective. Thus $F(g_1) \neq F(g_2)$, which shows that F is injective since g_1 and g_2 were arbitrary. $\square$

Lemma 7.7.2. *For sets A, B, and C, the set $(A^B)^C$ has the same cardinality as the set $A^{B\times C}$.*

Proof. We construct a bijection $F : A^{B\times C} \to (A^B)^C$. So, for any $f \in A^{B\times C}$, we have that $f : B \times C \to A$. Define $g : C \to A^B$ by $g(c) = h$ for any $c \in C$, where $h : B \to A$ is defined by $h(b) = f(b, c)$. Then assign $F(f) = g$.

To show that F is injective, consider $f, f' \in A^{B\times C}$ where $f \neq f'$. Then there is a $(b, c) \in B \times C$ where $f(b, c) \neq f'(b, c)$. Also let $g = F(f)$, $g' = F(f')$, $h = g(c)$, and $h' = g'(c)$. Then, by definition, we have $h(b) = f(b, c) \neq f'(b, c) = h'(b)$ so that $g(c) = h \neq h' = g'(c)$. From this it follows that $F(f) = g \neq g' = F(f')$, which shows that F is injective since f and f' were arbitrary.

Now consider any $g \in (A^B)^C$ and any $(b, c) \in B \times C$. Let $h = g(c) \in A^B$, and then assign $f(b, c) = h(b)$. Clearly then $f : B \times C \to A$ so that $f \in A^{B\times C}$. So let $g' = F(f)$ and consider any $c \in C$. Let $h = g(c)$ and $h' = g'(c)$ so that $h'(b) = f(b, c)$ by the definition of F. Consider any $b \in B$ so that $h(b) = f(b, c) = h'(b)$ by the definition of f. Since b was arbitrary, this shows that $g(c) = h = h' = g'(c)$. Since c was also arbitrary, this shows that $F(f) = g' = g$. Lastly, since g was arbitrary, this shows that F is surjective. $\square$

Main Problem.

Recall that we have $D = {\mathbb{Z}_+}^\omega = {\mathbb{Z}_+}^{\mathbb{Z}_+}$ and $E = X^\omega = X^{\mathbb{Z}_+}$, where we let $X = \{0, 1\}$. We show that these have the same cardinality.

Proof. First consider any $f \in E = X^{\mathbb{Z}_+}$. Then define $g(n) = f(n) + 1$ for $n \in \mathbb{Z}_+$ so that clearly $g \in {\mathbb{Z}_+}^{\mathbb{Z}_+} = D$. Now define the function $h : E \to D$ by $h(f) = g$. It is then trivial to show that h is an injection.

Now, for $n \in \mathbb{Z}_+$, define $x_n = 1$ and $x_i = 0$ when $i \in \mathbb{Z}_+$ and $i \neq n$. Clearly then $\mathbf{x} = (x_1, x_2, \ldots) \in X^{\mathbb{Z}_+}$, and it is easily shown that the function defined by $f(n) = \mathbf{x}$ is an injection from $\mathbb{Z}_+$ to $X^{\mathbb{Z}_+}$ Also clearly the identity function on $\mathbb{Z}_+$ is an injection since it is a bijection. It then follows from Lemma 7.7.1 that there is an injection $f_1 : {\mathbb{Z}_+}^{\mathbb{Z}_+} \to (X^{\mathbb{Z}_+})^{\mathbb{Z}_+}$, noting that clearly $X^{\mathbb{Z}_+} \neq \varnothing$.

We presently have that there is a bijection $f_2 : (X^{\mathbb{Z}_+})^{\mathbb{Z}_+} \to X^{\mathbb{Z}_+\times\mathbb{Z}_+}$ by Lemma 7.7.2, which is of course also an injection. Finally, since $\mathbb{Z}_+ \times \mathbb{Z}_+$ has the same cardinality as $\mathbb{Z}_+$ (by Corollary 7.4), it follows that there is an injection from $\mathbb{Z}_+ \times \mathbb{Z}_+$ to $\mathbb{Z}_+$. Since also the identity function on X is an injection, we have again that there is an injection $f_3 : X^{\mathbb{Z}_+\times\mathbb{Z}_+} \to X^{\mathbb{Z}_+}$ by Lemma 7.7.1. Thus clearly then $f_3 \circ f_2 \circ f_1$ is an injection from ${\mathbb{Z}_+}^{\mathbb{Z}_+} = D$ to $X^{\mathbb{Z}_+} = E$.

Therefore, since there is an injection from E to D as well as from D to E, it follows from Exercise 7.6b that D and E have the same cardinality as desired. $\square$

Exercise 7.8

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:

Again let A^B denote the set of functions from A to B.

Proof. First, for $\mathbf{x} \in X^\omega$, clearly the function that maps $\mathbf{x}$ to the set $\{\mathbf{x}\}$ is an injective function from X^ω to $\mathcal{B}$.

Now we construct an injection $f_1 : \mathcal{B} \to (X^\omega)^\omega$. So consider any $S \in \mathcal{B}$ so that S is a countable subset of X^ω. Then, by Theorem 7.1, we can choose a surjective function $g : \mathbb{Z}_+ \to S$. Note that this does require the Axiom of Choice since we must choose such a surjection for each $S \in \mathcal{B}$, and clearly $\mathcal{B}$ is infinite. Since $S \subset X^\omega$, g can be considered as a function from $\mathbb{Z}_+$ to X^ω so that $g \in (X^\omega)^\omega$, though of course it would no longer necessarily be surjective with this range. So we simply set $f_1(S) = g$

To show that f_1 is injective consider $S, S' \in \mathcal{B}$ where $S \neq S'$. Then, setting $g = f_1(S)$ and $g' = f_1(S')$, we have that $g(\mathbb{Z}_+) = S$ and $g'(\mathbb{Z}_+) = S'$ by definition. Since $S \neq S'$, we have that g and g' have the same domain but different image sets. Clearly this means that $f_1(S) = g \neq g' = f_1(S')$, which shows that f_1 is injective since S and S' were arbitrary.

Hence f_1 is an injection from $\mathcal{B}$ to $(X^\omega)^\omega = (X^{\mathbb{Z}_+})^{\mathbb{Z}_+}$. Now, from Lemma 7.7.2, we have that $(X^{\mathbb{Z}_+})^{\mathbb{Z}_+}$ has the same cardinality as $X^{\mathbb{Z}_+ \times \mathbb{Z}_+}$ so that there is a bijection $f_2 : (X^{\mathbb{Z}_+})^{\mathbb{Z}_+} \to X^{\mathbb{Z}_+ \times \mathbb{Z}_+}$, which is of course also an injection. Finally, since $\mathbb{Z}_+ \times \mathbb{Z}_+$ has the same cardinality as $\mathbb{Z}_+$ (by Corollary 7.4), it follows that there is an injection from $\mathbb{Z}_+ \times \mathbb{Z}_+$ to $\mathbb{Z}_+$. Since also the identity function on X is an injection, we have that there is an injection $f_3 : X^{\mathbb{Z}_+ \times \mathbb{Z}_+} \to X^{\mathbb{Z}_+}$ by Lemma 7.7.1. Then clearly $f_3 \circ f_2 \circ f_1$ is an injection from $\mathcal{B}$ to $X^{\mathbb{Z}_+} = X^\omega$.

Since there is an injection from X^ω to $\mathcal{B}$ and vice-versa, it follows that they have the same cardinality by Exercise 7.6b as desired. $\square$

Exercise 7.9

(a) The formula

$$(*) \qquad \begin{aligned} h(1) &= 1\,, \\ h(2) &= 2\,, \\ h(n) &= [h(n+1)]^2 - [h(n-1)]^2 \qquad \text{for } n \geq 2 \end{aligned}$$

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. [Hint: Reformulate $(*)$ so that the principle will apply and require h to be positive.]

(b) Show that the formula $(*)$ of part (a) does not determine h uniquely. [Hint: If h is a positive function satisfying $(*)$, let $f(i) = h(i)$ for $i \neq 3$, and let $f(3) = -h(3)$.]

(c) Show that there is no function $h : \mathbb{Z}_+ \to \mathbb{R}$ satisfying the formula

$$\begin{aligned} h(1) &= 1\,, \\ h(2) &= 2\,, \\ h(n) &= [h(n+1)]^2 + [h(n-1)]^2 \qquad \text{for } n \geq 2. \end{aligned}$$

Solution:

(a) First, notice that $(*)$ does not satisfy the principle of recursive definition because, for $n \geq 2$, $h(n)$ is not defined strictly in terms of values of h for positive integers less than n, since its definition depends on $h(n+1)$. Now we show that there does exists a function satisfying $(*)$.

Proof. Consider the following reformulation:

$$
\begin{aligned}
h(1) &= 1\,, \\
h(2) &= 2\,, \\
h(n) &= \sqrt{h(n-1) + [h(n-2)]^2} \qquad \text{for } n > 2\,,
\end{aligned}
$$

where as is convention we take the positive square root for $h(n)$. Clearly for $n \in \{1, 2\}$ we have that $h(n)$ is positive. Now suppose $n > 2$ and that $h(k)$ is positive for $k < n$ so that $h(n-1)$ and $h(n-2)$ are both positive. Then clearly $h(n-1) + [h(n-2)]^2$ is positive so that $h(n) = \sqrt{h(n-1) + [h(n-2)]^2}$ is defined and is positive. Hence $h(n)$ is positive and well-defined for all $n \in \mathbb{Z}_+$ by induction.

Thus, since $h(n)$ depends only on values of h for integers less than n, this satisfies the recursion principle so that a unique h satisfying the above exists. We also claim that this h satisfies $(*)$. Clearly the explicitly defined values of $h(1)$ and $h(2)$ are satisfied. For $n \geq 2$, we have that $n+1 > 2$ so that, by definition,

$$
\begin{aligned}
h(n+1) = \sqrt{h((n+1)-1) + [h((n+1)-2)]^2} &= \sqrt{h(n) + [h(n-1)]^2} \\
[h(n+1)]^2 &= h(n) + [h(n-1)]^2 \\
h(n) &= [h(n+1)]^2 - [h(n-1)]^2\,,
\end{aligned}
$$

which is the final constraint of $(*)$ so that it is also satisfied since $n \geq 2$ was arbitrary. □

(b) First note that, for the recursively defined function h from part (a),

$$
\begin{aligned}
h(3) &= \sqrt{h(2) + [h(1)]^2} = \sqrt{2 + 1^2} = \sqrt{3} \\
h(4) &= \sqrt{h(3) + [h(2)]^2} = \sqrt{\sqrt{3} + 2^2} = \sqrt{\sqrt{3} + 4}\,.
\end{aligned}
$$

Now define the function f as in the hint, that is $f(i) = h(i)$ for $i \neq 3$ and $f(3) = -h(3)$. Then we clearly have $f(3) = -h(3) = -\sqrt{3}$ while

$$
[f(4)]^2 - [f(2)]^2 = [h(4)]^2 - [h(2)]^2 = \left(\sqrt{\sqrt{3} + 4}\right)^2 - 2^2 = \sqrt{3} + 4 - 4 = \sqrt{3}
$$

so that $f(n) = -\sqrt{3} \neq \sqrt{3} = [f(n+1)]^2 - [f(n-1)]^2$ for $n = 3$, and hence $(*)$ is violated. So it would seem that the hint as given does not exactly work.

Now we show that the function satisfying $(*)$ is not unique, taking inspiration from the hint.

Proof. We construct a function f, different from h from part (a), that also satisfies $(*)$. We define f using recursion:

$$
\begin{aligned}
f(1) &= 1\,, \\
f(2) &= 2\,, \\
f(3) &= -\sqrt{3}\,, \\
f(n) &= \sqrt{f(n-1) + [f(n-2)]^2} \qquad \text{for } n > 3\,.
\end{aligned}
$$

Clearly, since each $f(n)$ is defined only in terms of $f(k)$ for $k < n$ (or without dependence on any values of f), f exists uniquely by the recursion principle so long as each $f(n)$ is well-defined. We show this presently by induction.

Clearly $f(n)$ is defined for $n \in \{1,2,3\}$. For $n = 4$ we have $f(n) = f(4) = \sqrt{f(3) + [f(2)]^2} = \sqrt{-\sqrt{3} + 2^2}\sqrt{-\sqrt{3} + 4}$. Now, since $1 < 3$, we have that $\sqrt{3} < 3 < 4$ so that $-\sqrt{3} + 4 = 4 - \sqrt{3} > 0$ and hence the square root, and therefore $f(4)$, is defined and positive. Now consider any $n > 4$ and suppose that $f(n-1)$ is positive. Then clearly $f(n) = \sqrt{f(n-1) + [f(n-2)]^2}$ is defined and positive since $f(n-1) > 0$, noting that even if $f(n-2) \leq 0$, its square is non-negative.. This completes the induction that shows that f is uniquely defined.

Clearly $f \neq h$ since $f(3) = -\sqrt{3} \neq \sqrt{3} = h(3)$. Also obviously $f(n)$ satisfies $(*)$ explicitly for $n \in \{1, 2\}$. For $n = 2$ we have

$$[f(n+1)]^2 - [f(n-1)]^2 = [f(3)]^2 - [f(1)]^2 = [-\sqrt{3}]^2 - 1^2 = 3 - 1 = 2 = f(n)\,.$$

Then, for $n > 2$ we have $n + 1 > 3$ so that, by definition,

$$\begin{aligned} f(n+1) = \sqrt{f((n+1)-1) + [f((n+1)-2)]^2} &= \sqrt{f(n) + [f(n-1)]^2} \\ [f(n+1)]^2 &= f(n) + [f(n-1)]^2 \\ f(n) &= [f(n+1)]^2 - [f(n-1)]^2\,. \end{aligned}$$

Thus the recursive part of $(*)$ holds for $n \geq 2$ so that $(*)$ holds over the whole domain of f as desired. □

(c)

Proof. Suppose that such a function h does exist. Since the recursive property holds for $n \geq 2$, we have

$$\begin{aligned} h(2) &= [h(3)]^2 + [h(1)]^2 \\ 2 &= [h(3)]^2 + 1^2 \\ [h(3)]^2 &= 2 - 1^2 = 1 \\ h(3) &= \pm 1\,. \end{aligned}$$

Similarly, we have

$$\begin{aligned} h(3) &= [h(4)]^2 + [h(2)]^2 \\ \pm 1 &= [h(4)]^2 + 2^2 \\ [h(4)]^2 &= \pm 1 - 2^2 = \pm 1 - 4 \end{aligned}$$

so that either $[h(4)]^2 = 1 - 4 = -3$ or $[h(4)]^2 = -1 - 4 = -5$. In either case we have $[h(4)]^2 < 0$, which is of course impossible since the square of a real number is always non-negative! So it must be that such a function does not exist. □

§8 The Principle of Recursive Definition

Exercise 8.1

Let $(b_1, b_2, \ldots)$ be an infinite sequence of real numbers. The sum $\sum_{k=1}^{n} b_k$ is defined by induction as follows:

$$\sum_{k=1}^{n} b_k = b_1 \qquad \text{for } n = 1,$$

$$\sum_{k=1}^{n} b_k = \left(\sum_{k=1}^{n-1} b_k\right) + b_n \qquad \text{for } n > 1.$$

Let A be the set of real numbers; choose ρ so that Theorem 8.4 applies to define this sum rigorously. We sometimes denote the sum $\sum_{k=1}^{n} b_k$ by the symbol $b_1 + b_2 + \cdots + b_n$.

Solution:

For a function $f : \{1, \ldots, m\} \to A$, define $\rho(f) = f(m) + b_{m+1}$. For clarity, denote the sum function by $s : \mathbb{Z}_+ \to A$ so that $s(n) = \sum_{k=1}^{n} b_k$. Then by Theorem 8.4 there is a unique $s : \mathbb{Z}_+ \to A$ such that

$$\begin{aligned} s(1) &= b_1\,, \\ s(n) &= \rho(s \restriction \{1, \ldots, n-1\}) \qquad \text{for } n > 1. \end{aligned}$$

Then we clearly have that $\sum_{k=1}^{1} b_k = s(1) = b_1$ and

$$\sum_{k=1}^{n} b_k = s(n) = \rho(s \restriction \{1, \ldots, n-1\}) = s(n-1) + b_{(n-1)+1} = \sum_{k=1}^{n-1} b_k + b_n$$

for $n > 1$ as desired.

Exercise 8.2

Let $(b_1, b_2, \ldots)$ be an infinite sequence of real numbers. We define the product $\prod_{k=1}^{n} b_k$ by the equations

$$\begin{aligned} \prod_{k=1}^{1} b_k &= b_1\,, \\ \prod_{k=1}^{n} b_k &= \left(\prod_{k=1}^{n-1} b_k\right) \cdot b_n \qquad \text{for } n > 1. \end{aligned}$$

Use Theorem 8.4 to define the product rigorously. We sometimes denote the product $\prod_{k=1}^{n} b_k$ by the symbol $b_1 b_2 \cdots b_n$.

Solution:

First, for any function $f : \{1, \ldots, m\} \to \mathbb{R}$, define ρ by $\rho(f) = f(m) \cdot b_{m+1}$. Then, by the recursion theorem (Theorem 8.4), there is a unique function $p : \mathbb{Z}_+ \to \mathbb{R}$ such that

$$\begin{aligned} p(1) &= b_1\,, \\ p(n) &= \rho(p \restriction \{1, \ldots, n-1\}) \qquad \text{for } n > 1. \end{aligned}$$

Then we define $\prod_{k=1}^{n} b_k = p(n)$ so that we have $\prod_{k=1}^{1} b_k = p(1) = b_1$ and

$$\prod_{k=1}^{n} b_k = p(n) = \rho(p \restriction \{1, \ldots, n-1\}) = p(n-1) \cdot b_{(n-1)+1} = \left(\prod_{k=1}^{n-1} b_k\right) \cdot b_n$$

for $n > 1$ as desired.

Exercise 8.3

Obtain the definitions of a^n and $n!$ for $n \in \mathbb{Z}_+$ as special cases of Exercise 8.2.

Solution:

Regarding a^n, defined the sequence $(b_1, b_2, \ldots)$ by $b_i = a$ for every $i \in \mathbb{Z}_+$, which we could denote by $(a, a, \ldots)$. Then define $a^n = \prod_{k=1}^n b_k$ as it is defined in Exercise 8.2, and we claim that this satisfies the inductive definition given in Exercise 4.6 and Example 8.2.

Proof. First, we clearly have $a^1 = \prod_{k=1}^1 b_k = b_1 = a$. Next, for $n > 1$, we have

$$*a^n = \prod_{k=1}^n b_k = \left(\prod_{k=1}^{n-1} b_k\right) \cdot b_n = a^{n-1} \cdot a\,,$$

which shows that the inductive definition is satisfied. □

Since it does not seem to be given in the book thus far, we reiterate the typical inductive definition for $n!$:

$$\begin{aligned} 1! &= 1\,, \\ n! &= (n-1)! \cdot n \qquad \text{for } n > 1. \end{aligned}$$

Now, define the sequence $(b_1, b_2, \ldots)$ by $b_i = i$ for $i \in \mathbb{Z}_+$. We then claim that defining $n! = \prod_{k=1}^n b_k$ as defined in Exercise 8.2 satisfies this definition.

Proof. First, we have $1! = \prod_{k=1}^1 b_k = b_1 = 1$. Then we also have

$$n! = \prod_{k=1}^n b_k = \left(\prod_{k=1}^{n-1} b_k\right) \cdot b_n = (n-1)! \cdot n$$

for $n > 1$ so that the definition is clearly satisfied. □

Exercise 8.4

The *Fibonacci numbers* of number theory are defined recursively by the formula

$$\begin{aligned} \lambda_1 &= \lambda_2 = 1\,, \\ \lambda_n &= \lambda_{n-1} + \lambda_{n-2} \qquad \text{for } n > 2. \end{aligned}$$

Define them rigorously by use of Theorem 8.4.

Solution:

First, note that the Fibonacci numbers are all positive integers. So, for any function $f : \{1, \ldots, m\} \to \mathbb{Z}_+$ define

$$\rho(f) = \begin{cases} 1 & m = 1 \\ f(m) + f(m-1) & m > 1\,, \end{cases}$$

noting that clearly the range of ρ is still $\mathbb{Z}_+$ since that is the range of f. Then, by Theorem 8.4, there is a unique function $F : \mathbb{Z}_+ \to \mathbb{Z}_+$ such that

$$F(1) = 1\,,$$

$$F(n) = \rho(F \restriction \{1, \ldots, n-1\}) \qquad \text{for } n > 1.$$

We claim that the Fibonacci numbers are $\lambda_n = F(n)$ for $n \in \mathbb{Z}_+$.

Proof. To show that the numbers λ_n satisfy the inductive definition of the Fibonacci numbers, first note that we clearly have $\lambda_1 = F(1) = 1$. We also have that

$$\lambda_2 = F(2) = \rho(F \restriction \{1\}) = 1\,.$$

Lastly, for any $n > 2$, clearly $n > 1$ also and $n - 1 > 1$ so that

$$\lambda_n = F(n) = \rho(F \restriction \{1, \ldots, n-1\}) = F(n-1) + F([n-1]-1) = \lambda_{n-1} + \lambda_{n-2}\,,$$

which shows that the inductive definition is satisfied. $\square$

Exercise 8.5

Show that there is a unique function $h : \mathbb{Z}_+ \to \mathbb{R}_+$ satisfying the formula

$$\begin{aligned} h(1) &= 3\,, \\ h(i) &= [h(i-1)+1]^{1/2} \qquad \text{for } i > 1. \end{aligned}$$

Solution:

Proof. First, for any function $f : \{1, \ldots, m\} \to \mathbb{R}_+$, define

$$\rho(f) = [f(m)+1]^{1/2}\,.$$

Consider any $m \in \mathbb{Z}_+$ and any function $f : \{1, \ldots, m\} \to \mathbb{R}_+$. Since $f(m) \in \mathbb{R}_+$, it follows that $f(m) + 1 \in \mathbb{R}_+$ also so that $\rho(f) = [f(m)+1]^{1/2}$ is defined and is positive. Hence ρ is a well-defined function with range $\mathbb{R}_+$ since m and f were arbitrary. It then follows from the principle of recursive definition (Theorem 8.4) that there is a unique function $h : \mathbb{Z}_+ \to \mathbb{R}_+$ such that

$$\begin{aligned} h(1) &= 3\,, \\ h(n) &= \rho(h \restriction \{1, \ldots, n-1\}) \qquad \text{for } n > 1. \end{aligned}$$

It is easy to see that this h satisfies the required property since $h(1) = 3$ and

$$h(i) = \rho(h \restriction \{1, \ldots, i-1\}) = [h(i-1)+1]^{1/2}$$

for $i > 1$ as desired.

Now we show that such a function is unique. Suppose that g and h both satisfy the inductive formula. We show by induction that $g(i) = h(i)$ for all $i \in \mathbb{Z}_+$, from which it clearly follows that $g = h$. First, we clearly have $g(1) = 3 = h(1)$. Now suppose that $g(i) = h(i)$ for $i \in \mathbb{Z}_+$. Then we have that $i + 1 > 1$ so that $g(i+1) = [g(i)+1]^{1/2} = [h(i)+1]^{1/2} = h(i+1)$ since $g(i) = h(i)$ and we are taking the positive root. This completes the induction. $\square$

Exercise 8.6

(a) Show that there is no function $h : \mathbb{Z}_+ \to \mathbb{R}_+$ satisfying the formula

$$\begin{aligned} h(1) &= 3\,, \\ h(i) &= [h(i-1) - 1]^{1/2} \qquad \text{for } i > 1. \end{aligned}$$

Explain why this example does not violate the principle of recursive definition.

(b) Consider the recursion formula

$$\begin{aligned} h(1) &= 3\,, \\ h(i) &= \begin{cases} [h(i-1) - 1]^{1/2} & \text{if } h(i-1) > 1 \\ 5 & \text{if } h(i-1) \le 1 \end{cases} \qquad \text{for } i > 1. \end{aligned}$$

Show that there exists a unique function $h : \mathbb{Z}_+ \to \mathbb{R}_+$ satisfying this formula.

Solution:

(a)

Proof. Suppose to the contrary that there is such a function h. Then clearly $h(1) = 3$ and $h(2) = \sqrt{h(1) - 1} = \sqrt{3 - 1} = \sqrt{2}$. Now, since $1 < 2 < 4$, we clearly have $1 < \sqrt{2} < \sqrt{4} = 2$. Thus $0 < \sqrt{2} - 1 < 1$ so that $h(3) = \sqrt{h(2) - 1} = \sqrt{\sqrt{2} - 1}$ is defined. However, we also have that $0 < h(3) = \sqrt{\sqrt{2} - 1} < 1$ since $0 < \sqrt{2} - 1 < 1$, and hence $h(3) - 1 < 0$. We then have that

$$\begin{aligned} h(4) &= \sqrt{h(3) - 1} \\ [h(4)]^2 &= h(3) - 1 < 0\,, \end{aligned}$$

which is of course impossible since a square is always non-negative. This contradiction shows that such a function h cannot exist. □

Note that this does not ostensibly violate the principle of recursive definition since $h(n)$ is defined only in terms of values of h less than n for $n > 1$. However, were one to try to show the existence of h rigorously using the principle, one would find that the required function ρ would not be well-defined.

(b)

Proof. First, for any function $f : \{1, \ldots, m\} \to \mathbb{R}_+$, define

$$\rho(f) = \begin{cases} [f(m) - 1]^{1/2} & f(m) > 1 \\ 5 & f(m) \le 1\,. \end{cases}$$

Consider any $m \in \mathbb{Z}_+$ and any function $f : \{1, \ldots, m\} \to \mathbb{R}_+$. If $f(m) > 1$ then clearly $f(m) - 1 > 0$ so that $\rho(f) = [f(m) - 1]^{1/2}$ is defined and positive. If $f(m) \le 1$ then clearly $\rho(f) = 5$ is also defined and positive. Since m and f were arbitrary, this shows that ρ is a well-defined function with range $\mathbb{R}_+$.

It then follows from the principle of recursive definition (Theorem 8.4) that there is a unique function $h : \mathbb{Z}_+ \to \mathbb{R}_+$ such that

$$\begin{aligned} h(1) &= 3\,, \\ h(n) &= \rho(h \restriction \{1, \ldots, n-1\}) \qquad \text{for } n > 1. \end{aligned}$$

To see that this h satisfies the recursion formula, clearly $h(1) = 3$, and, for $i > 1$, we have

$$h(i) = \rho(h \restriction \{1, \ldots, i-1\}) = \begin{cases} [h(i-1) - 1]^{1/2} & h(i-1) > 1 \\ 5 & h(i-1) \leq 1 \end{cases}$$

as desired.

To show that this function is unique, suppose that g and h both satisfy the recursive formula. We show by induction that $g(n) = h(n)$ for all $n \in \mathbb{Z}_+$ so that clearly $g = h$. First, obviously $g(1) = 3 = h(1)$. Now suppose that $g(n) = h(n)$ for $n \in \mathbb{Z}_+$ so that $n + 1 > 1$. Then, if $g(n) = h(n) > 1$ then we have $g(n+1) = [g(n) - 1]^{1/2} = [h(n) - 1]^{1/2} = h(n+1)$ since $g(n) = h(n)$ and the roots are taken to be positive. Similarly, if $g(n) = h(n) \leq 1$, then $g(n+1) = 5 = h(n+1)$. Thus in either case $g(n+1) = h(n+1)$, which completes the induction. $\square$

Exercise 8.7

Prove Theorem 8.4.

Solution:

The proof follows the same pattern used to prove $(*)$ at the beginning of the section, which culminates in Theorem 8.3. Similar to that approach, two lemmas will be proved first. In what follows, $(*)$ refers to the properties defined in the statement of Theorem 8.4.

Lemma 8.7.1. *Given $n \in \mathbb{Z}_+$, there exists a function $f : \{1, \ldots, n\} \to A$ that satisfies $(*)$ for all i in its domain.*

Proof. We show this by induction on n. First, for $n = 1$, clearly the function $f : \{1\} \to A$ defined by $f(1) = a_0$ satisfies $(*)$. Now suppose that $(*)$ holds for some function $f' : \{1, \ldots, n\} \to A$ for $n \in \mathbb{Z}_+$. Now define $f : \{1, \ldots, n+1\} \to A$ by

$$f(i) = \begin{cases} f'(i) & i \in \{1, \ldots, n\} \\ \rho(f') & i = n+1 \end{cases}$$

for any $i \in \{1, \ldots, n+1\}$. Note that f is not defined in terms of itself, but in terms of f' and ρ.

First, we clearly have $f' = f \restriction \{1, \ldots, n\}$ since $f(i) = f'(i)$ for all $i \in \{1, \ldots, n\}$. Then, clearly $f(1) = f'(1) = a_0$ since $1 \leq n$ and f' satisfies $(*)$. Consider any $i \in \{1, \ldots, n+1\}$ where $i > 1$. Then we have

$$f(i) = f'(i) = \rho(f' \restriction \{1, \ldots, i-1\}) = \rho(f \restriction \{1, \ldots, i-1\})$$

if $1 < i \leq n$ since f' satisfies $(*)$. Lastly, if $i = n+1$, then

$$f(i) = \rho(f') = \rho(f \restriction \{1, \ldots, n\}) = \rho(f \restriction \{1, \ldots, i-1\})$$

again. This shows that f satisfies $(*)$, thereby completing the induction. $\square$

Lemma 8.7.2. *Suppose that $f : \{1, \ldots, n\} \to A$ and $g : \{1, \ldots, m\} \to C$ both satisfy $(*)$ for all i in their respective domains. Then $f(i) = g(i)$ for all i in both domains.*

Proof. Suppose that this is not the case and let i be the *smallest* integer (in the domain of both f and g) for which $f(i) \neq g(i)$. Hence $f(j) = g(j)$ for all $1 \leq j < i$ so that clearly $f \restriction \{1, \ldots, i-1\} = g \restriction \{1, \ldots, i-1\}$. Now, it cannot be that $i = 1$ since clearly $f(1) = a_0 = g(1)$. So then it must be that $1 < i$ so that $f(i) = \rho(f \restriction \{1, \ldots, i-1\})$ and $g(i) = \rho(g \restriction \{1, \ldots, i-1\})$ since they both satisfy $(*)$. Since $f \restriction \{1, \ldots, i-1\} = g \restriction \{1, \ldots, i-1\}$, we then clearly have

$$f(i) = \rho(f \restriction \{1, \ldots, i-1\}) = \rho(g \restriction \{1, \ldots, i-1\}) = g(i)$$

in contradiction with the definition of i. Thus the result must be true as desired. $\square$

Main Problem.

Proof. Lemmas 8.7.1 and 8.7.2 show that there exists a unique function $f_n : \{1, \ldots, n\} \to A$ satisfying $(*)$ for every $n \in \mathbb{Z}_+$. We then define $h = \bigcup_{n \in \mathbb{Z}_+} f_n$ and claim that this is the unique function from $\mathbb{Z}_+$ to A satisfying $(*)$.

First we must show that h is a function at all. So consider any $i \in \mathbb{Z}_+$ and suppose that (i, x) and (i, y) are in h. Then there are $n, m \in \mathbb{Z}_+$ where $(i, x) \in f_n$ and $(i, y) \in f_m$ since $h = \bigcup_{n \in \mathbb{Z}_+} f_n$, noting that it must be that $i \leq n$ and $i \leq m$. Since f_n and f_m both satisfy $(*)$ and clearly i is in the domain of both, it follows from Lemma 8.7.2 that $x = f_n(i) = f_m(i) = y$. This shows that h is a function since (i, x) and (i, y) were arbitrary. Also, clearly the domain of h is $\mathbb{Z}_+$ since, for any $i \in \mathbb{Z}_+$, i is in the domain of f_i and so in the domain of h. Lastly, clearly the range of h is A since that is the range of all the f_n functions.

Now we show that h satisfies $(*)$. First we have that 1 is clearly in the domain of h and f_1 so that it has to be that $h(1) = f_1(1) = a_0$ since h is a function, $f_1 \subset h$, and f_1 satisfies $(*)$. Now suppose that $i > 1$. Then clearly i is in the domain of h and f_i so that it has to be that $h(j) = f_i(j)$ for $1 \leq j \leq i$ since h was shown to be a function and $f_i \subset h$. It then follows that $h \restriction \{1, \ldots, i-1\} = f_i \restriction \{1, \ldots, i-1\}$. Thus we have

$$h(i) = f_i(i) = \rho(f_i \restriction \{1, \ldots, i-1\}) = \rho(h \restriction \{1, \ldots, i-1\})$$

since f_i satisfies $(*)$. This completes the proof that h also satisfies $(*)$.

Lastly, we show that h is unique, which is very similar to the proof of Lemma 8.7.2. So suppose that f and g are two functions from $\mathbb{Z}_+$ to A that both satisfy $(*)$. Suppose also that $f \neq g$ so that there is a smallest integer i such that $f(i) \neq g(i)$. Now, it cannot be that $i = 1$ since we have $f(1) = a_0 = g(1)$ since they both satisfy $(*)$. Hence $i > 1$ and, since i is the smallest integer where $f(i) \neq g(i)$, it follows that $f(j) = g(j)$ for all $1 \leq j < i$. Therefore we have that $f \restriction \{1, \ldots, i-1\} = g \restriction \{1, \ldots, i-1\}$ so that

$$f(i) = \rho(f \restriction \{1, \ldots, i-1\}) = \rho(g \restriction \{1, \ldots, i-1\}) = g(i)$$

since f and g both satisfy $(*)$ and $i > 1$. This of course contradicts the definition of i so that it has to be that in fact $f = g$. This shows the uniqueness of h constructed above. $\square$

Exercise 8.8

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, an element $\rho(f)$ of A. Then there is a unique function $h : \mathbb{Z}_+ \to A$ such that $h(n) = \rho(h \restriction S_n)$ for each $n \in \mathbb{Z}_+$.

Solution:

Denote the above property of h by $(*)$. We show that there is a unique $h : \mathbb{Z}_+ \to A$ satisfying this using the standard principle of recursive definition, Theorem 8.4.

Proof. First, note that $S_1 = \{n \in \mathbb{Z}_+ \mid n < 1\} = \varnothing$ by definition. Note also that $\varnothing$ itself is vacuously a function from $S_1 = \varnothing$ to A, and is the only such function. It then follows that $f \restriction S_1 = f \restriction \varnothing = \varnothing$ for any $f : S_n \to A$ for some $n \in \mathbb{Z}_+$. So then, define $a_0 = \rho(\varnothing)$ so that there is a unique function h such that

$$\begin{aligned} h(1) &= a_0\,, \\ h(i) &= \rho(h \restriction \{1, \ldots, i-1\}) \qquad \text{for } i > 1. \end{aligned}$$

by Theorem 8.4. Denote this property by $(+)$.

We first claim that this h satisfies $(*)$. To see this, consider any $n \in \mathbb{Z}_+$. If $n = 1$ then we have

$$h(n) = h(1) = a_0 = \rho(\varnothing) = \rho(h \restriction \varnothing) = \rho(h \restriction S_1) = \rho(h \restriction S_n)\,.$$

If $n > 1$ then by $(+)$ we have

$$h(n) = \rho(h \restriction \{1, \ldots, n-1\}) = \rho(h \restriction S_n)$$

again. Since n was arbitrary, this shows that $(*)$ is satisfied.

To show that this h satisfying $(*)$ is unique, suppose that another function $f : \mathbb{Z}_+ \to A$ satisfies $(*)$. Then we have

$$h(1) = \rho(h \restriction S_1) = \rho(h \restriction \varnothing) = \rho(\varnothing) = a_0$$

and

$$h(i) = \rho(h \restriction S_i) = \rho(h \restriction \{1, \ldots, i-1\})$$

for $i > 1$. This shows that f also satisfies $(+)$, and, since we know that the function satisfying $(+)$ is unique, it must be that $f = h$ as desired. $\square$

§9 Infinite Sets and the Axiom of Choice

Exercise 9.1

Define an injective map $f : \mathbb{Z}_+ \to X^\omega$, where X is the two-element set $\{0, 1\}$, without using the choice axiom.

Solution:

For any $n \in \mathbb{Z}_+$, define

$$x_i = \begin{cases} 0 & i \neq n \\ 1 & i = n \end{cases}$$

for $i \in \mathbb{Z}_+$. Then set $f(n) = \mathbf{x} = (x_1, x_2, \ldots)$ so that clearly f is a function from $\mathbb{Z}_+$ to X^ω. It is easy to show that f is injective.

Proof. Consider $n, m \in \mathbb{Z}_+$ where $n \neq m$. Then let $\mathbf{x} = f(n)$ and $\mathbf{y} = f(m)$. Then we have that $x_n = 1$ while $y_n = 0$ by the definition of f since $n \neq m$. It thus follows that $f(n) = \mathbf{x} \neq \mathbf{y} = f(m)$, which shows that f is injective since n and m were arbitrary. $\square$

Exercise 9.2

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:

Lemma 9.2.1. *If A is a countable set and $\mathcal{A}$ is the collection of nonempty subsets of A then $\mathcal{A}$ has a choice function.*

Proof. Since A is countable, there is an injective function $A \to \mathbb{Z}_+$ by Theorem 7.1. We define a choice function $c : \mathcal{A} \to \bigcup_{B\in\mathcal{A}} B$. Consider any $X \in \mathcal{A}$ so that X is a nonempty subset of A. Then $f(X)$ is a nonempty subset of $\mathbb{Z}_+$ so that it has a unique smallest element n since $\mathbb{Z}_+$ is well-ordered. Now, since $n \in f(X)$, clearly there is an $x \in X$ such that $f(x) = n$. Moreover, it follows from the fact that f is injective that this x is unique. So set $c(X) = x$ so that clearly x is a choice function on $\mathcal{A}$ since $c(X) = x \in X$. $\square$

Main Problem.

(a) Since $\mathbb{Z}_+$ is countable, a choice function can be constructed as in Lemma 9.2.1.

(b) Since $\mathbb{Z}$ is countable (by Example 7.1), a choice function can be constructed as in Lemma 9.2.1.

(c) Since $\mathbb{Q}$ is countable (by Exercise 7.1), a choice function can be constructed as in Lemma 9.2.1.

(d) First, there is an injective function f from the real interval $[0, 1]$ to X^ω. The most straightforward such function is, for each $x \in [0, 1]$ let $0.x_1x_2x_3\ldots$ be a unique binary expansion of x (these can be made unique by avoiding binary expansions that end in all 1's, noting though that the expansion of 1 itself must be $0.111\ldots$). So suppose that c were a choice function on $\mathcal{D}$ (that is presumably constructed without the choice axiom). If X is a nonempty subset of $[0, 1]$ then $f(X)$ is a set in $\mathcal{D}$ so that we can choose $c(f(X)) \in f(X)$. Since f is injective, there is a unique $x \in X$ where $f(x) = c(f(X))$, and so choosing x results in a choice function on the collection of nonempty subsets of $[0, 1]$ since X was arbitrary.

This would allow one to then well-order $[0, 1]$ without using the choice axiom, which evidently nobody has done. As far as I have been able to determine, this has not yet been proven impossible, it is just that nobody has been able to do it. So it would seem that such an explicit construction of a choice function on $\mathcal{D}$ would at least make one famous. Or else it is impossible, which is what we assume to be the case here.

Exercise 9.3

Suppose that A is a set and $\{f_n\}_{n\in\mathbb{Z}_+}$ is a given indexed family of injective functions

$$f_n : \{1, \ldots, n\} \to A\,.$$

Show that A is infinite. Can you define an injective function $f : \mathbb{Z}_+ \to A$ without using the choice axiom?

Solution:

We defer the proof that A is infinite until we define an injective $f : \mathbb{Z}_+ \to A$, which we can do without using the choice axiom by using the principle of recursive definition.

Proof. First, let $a_0 = f_1(1) \in A$. Now consider any function $g : S_n \to A$. If $g = \varnothing$ then set $\rho(\varnothing) = \rho(g) = a_0$. Otherwise let $I_g = \{i \in S_{n+1} \mid f_n(i) \notin g(S_n)\}$. Suppose for the moment that $I_g = \varnothing$. Consider any $x \in f_n(S_{n+1})$ so that there is a $k \in S_{n+1}$ where $f_n(k) = x$. Then it has to be that $x = f_n(k) \in g(S_n)$ since otherwise we would have $k \in I_g$. Since x was arbitrary, this shows that $f_n(S_{n+1}) \subset g(S_n)$. Thus the identity function $h_1 : f_n(S_{n+1}) \to g(S_n)$ is an injection. Clearly g is a surjection from S_n to its image $g(S_n)$ so that there is we can construct a particular injection $h_2 : g(S_n) \to S_n$ by Corollary 6.7. Lastly, f_n is an injection from S_{n+1} to $f_n(S_{n+1})$. Therefore $h = h_2 \circ h_1 \circ f_n$ is an injection from S_{n+1} to S_n. Hence h is a bijection from S_{n+1} to $h(S_{n+1})$, which is clearly a subset of S_n since S_n is the range of h. But, since $S_n \subsetneq S_{n+1}$, clearly $h(S_{n+1}) \subsetneq S_{n+1}$ as well so that h is a bijection from S_{n+1} onto a proper subset of itself. As S_{n+1} is clearly finite, this violates Corollary 6.3 so that we have a contradiction.

So it must be that $I_g \neq \varnothing$ so that it is a nonempty set of positive integers, and hence has a smallest element i. So simply set $\rho(g) = f_n(i)$. Now, it then follows from the principle of recursive definition that there is a unique $f : \mathbb{Z}_+ \to A$ such that

$$\begin{aligned} f(1) &= a_0\,, \\ f(n) &= \rho(f \restriction S_n) \qquad \text{for } n > 1. \end{aligned}$$

We claim that this f is injective.

To see this we first show that $f(n) \notin f(S_n)$ for all $n \in \mathbb{Z}_+$. If $n = 1$ we have that $f(n) = f(1) = a_0 = f_1(1)$ and $f(S_n) = f(\varnothing) = \varnothing$ so that clearly the result holds. If $n > 1$ then $f(n) = \rho(f \restriction S_n) = f_n(i)$ for some $i \in I_{f\restriction S_n}$ since clearly $S_n \neq \varnothing$ so that $f \restriction S_n \neq \varnothing$. Since $i \in I_{f\restriction S_n}$ we have that $f(n) = f_n(i) \notin (f \restriction S_n)(S_n) = f(S_n)$ as desired. This shows that f is injective. For consider any $n, m \in \mathbb{Z}_+$ where $n \neq m$. Without loss of generality we can assume that $n < m$. Then clearly $f(n) \in f(S_m)$ since $n \in S_m$ since $n < m$. However, by what was just shown, we have $f(m) \notin f(S_m)$ so that it has to be that $f(n) \neq f(m)$. This shows f to be injective since n and m were arbitrary.

Lastly, since $f : \mathbb{Z}_+ \to A$ is injective, it follows that f is a bijection from $\mathbb{Z}_+$ to $f(\mathbb{Z}_+) \subset A$. Hence $f(\mathbb{Z}_+)$ is infinite since $\mathbb{Z}_+$ is, and since it is a subset of A, it has to be that A is infinite as well. ☐

Exercise 9.4

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 use of the choice axiom. (Several of the earlier exercises have used the choice axiom also.)

Solution:

This was the proof of Theorem 7.5, which asserts that a countable union of countable sets is also countable. The following rewritten proof makes explicit use of the choice axiom and so points out where it is needed.

Proof. Let $\{A_n\}_{n\in J}$ be an indexed family of countable sets, where the index set J is $\{1, \ldots, N\}$ or $\mathbb{Z}_+$. Assume that each set A_n is nonempty for convenience since this does not change anything. Now, for each $n \in J$, let B_n be the set of surjective functions from $\mathbb{Z}_+$ to A_n. Since each A_n is countable, it follows from Theorem 7.1 that $B_n \neq \varnothing$. Then, by the axiom of choice, the collection $\{B_n\}_{n\in J}$ has a choice function c such that $c(B_n) \in B_n$ for every $n \in J$.

Now set $f_n = c(B_n)$ for every $n \in J$ so that $f_n \in B_n$ and hence is a surjection from $\mathbb{Z}_+$ into A_n. Since J is countable, there is also a surjection $g : \mathbb{Z}_+ \to J$ by Theorem 7.1. Then define $h : \mathbb{Z}_+ \times \mathbb{Z}_+ \to \bigcup_{n \in J} A_n$ by $h(k,m) = f_{g(k)}(m)$ for $k, m \in \mathbb{Z}_+$.

We now show that h is surjective. So consider any $a \in \bigcup_{n \in J} A_n$ so that $a \in A_n$ for some $n \in J$. Since $g : \mathbb{Z}_+ \to J$ is surjective, there is a $k \in \mathbb{Z}_+$ where $g(k) = n$. Also, since $f_n : \mathbb{Z}_+ \to A_n$ is surjective, there is an $m \in \mathbb{Z}_+$ where $f_n(m) = a$. We then have by definition that

$$h(k,m) = f_{g(k)}(m) = f_n(m) = a\,,$$

which shows that h is surjective since a was arbitrary.

Lastly, since $\mathbb{Z}_+ \times \mathbb{Z}_+$ is countable by Example 7.2, there is a bijection $h' : \mathbb{Z}_+ \to \mathbb{Z}_+ \times \mathbb{Z}_+$. It then follows that $h \circ h'$ is a surjection from $\mathbb{Z}_+$ to $\bigcup_{n \in J} A_n$, which shows that $\bigcup_{n \in J} A_n$ is countable again by Theorem 7.1. □

Exercise 9.5

(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:

(a)

Proof. Suppose that $f : A \to B$ is surjective. Now, by the choice axiom, the collection $\mathcal{A} = \mathcal{P}(A) - \{\varnothing\}$ is a collection of nonempty sets and thus has a choice function c. Consider any $b \in B$ and the set $A_b = \{x \in A \mid f(x) = b\}$. Then $A_b \neq \varnothing$ since f is surjective, and hence $A_b \in \mathcal{A}$ since clearly also $A_b \subset A$ so that $A_b \in \mathcal{P}(A)$. So set $h(b) = c(A_b) \in A_b$ so that $h(b) \in A$ since $A_b \subset A$. Hence h is a function from B to A.

Recall that, by definition, h is a right inverse if and only if $f \circ h = i_B$, which we show presently. So consider any $b \in B$ and let $a = h(b) = c(A_b) \in A_b$ so that $f(a) = b$. Then clearly

$$(f \circ h)(b) = f(h(b)) = f(a) = b\,,$$

which shows that $f \circ h = i_B$ since b was arbitrary. Hence h is a right inverse of f. □

(b)

Proof. Suppose that $f : A \to B$ is injective and $A \neq \varnothing$. Then f is a bijection from A to its image $f(A) \subset B$ and hence its inverse f^{-1} is a function from $f(A)$ to A. Now, since A is nonempty, there is an $a_0 \in A$. So define $h : B \to A$ by

$$h(b) = \begin{cases} f^{-1}(b) & b \in f(A) \\ a_0 & b \notin f(A) \end{cases}$$

for any $b \in B$. Recall that h is a left inverse of f if and only if $h \circ f = i_A$ by definition, which we show now.

So consider any $a \in A$ and let $b = f(a)$ so that clearly $b \in f(A)$. Hence by definition $h(b) = f^{-1}(b) = f^{-1}(f(a)) = a$. Finally, we have

$$(h \circ f)(a) = h(f(a)) = h(b) = a\,.$$

This shows that $h \circ f = i_A$ since a was arbitrary. Therefore h is a left inverse of f as desired. $\square$

Note that this proof does not require the axiom of choice as we did not need to make a choice for each $b \in B$ in order to define h as we did in part A.

Exercise 9.6

Most of the famous paradoxes of naive set theory are associated in some way or another 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 elements of themselves:

$$\mathcal{B} = \{A \mid A \in \mathcal{A} \text{ and } A \notin A\}\ .$$

(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:

(a) We show that $\mathcal{P}(\mathcal{A}) \subset \mathcal{A}$ and that a contradiction results.

Proof. Consider any set $A \in \mathcal{P}(\mathcal{A})$. Since A is a set and $\mathcal{A}$ is the set of all sets, clearly $A \in \mathcal{A}$ and hence $\mathcal{P}(\mathcal{A}) \subset \mathcal{A}$ since A was arbitrary. Therefore the identity function $i_{\mathcal{P}(\mathcal{A})}$ is clearly an injection from $\mathcal{P}(\mathcal{A})$ to $\mathcal{A}$. However, this is impossible by Theorem 7.8! Hence we have reached a contradiction. $\square$

(b) We show that the existence of $\mathcal{B}$ is a contradiction by showing that supposing either $\mathcal{B} \in \mathcal{B}$ or $\mathcal{B} \notin \mathcal{B}$ results in a contradiction.

Proof. Suppose that $\mathcal{B} \in \mathcal{B}$ so that by definition we have $\mathcal{B} \in \mathcal{A}$ and $\mathcal{B} \notin \mathcal{B}$, the latter of which clearly contradicts our initial supposition. On the other hand, suppose that $\mathcal{B} \notin \mathcal{B}$. Then, since clearly also $\mathcal{B} \in \mathcal{A}$ since it is a set, it follows that $\mathcal{B} \in \mathcal{B}$ by definition. This again contradicts the initial supposition. Since one or the other ($\mathcal{B} \in \mathcal{B}$ or $\mathcal{B} \notin \mathcal{B}$) must be true, we are then guaranteed to have a contradiction. $\square$

Exercise 9.7

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 B has greater cardinality than C, then A has greater cardinality than C.

(c) Find a sequence $A_1, A_2, \dots$ 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 cardinality greater than A_n.

Solution:

Lemma 9.7.1. *For any set A, $\mathcal{P}(A)$ has greater cardinality than A.*

Proof. Clearly the function that maps $a \in A$ to $\{a\} \in \mathcal{P}(A)$ is an injection. However, we know from Theorem 7.8 that there is no injection from $\mathcal{P}(A)$ to A. Together these show that $\mathcal{P}(A)$ has greater cardinality than A as desired. □

Main Problem.

(a)

Proof. Suppose that A is any uncountable set. Clearly A is not finite for then it would be countable. Hence it is infinite and so there is a injection from $\mathbb{Z}_+$ to A by Theorem 9.1. There also cannot be an injection from A to $\mathbb{Z}_+$, for if there were then A would be countable by Theorem 7.1. This shows that A has greater cardinality than $\mathbb{Z}_+$ by definition. □

(b)

Proof. Since A has greater cardinality than B, there is an injection $f : B \to A$. Likewise, since B has greater cardinality than C, there is an injection $g : C \to B$. It then follows that $f \circ g$ is an injection of C into A. Now suppose that $h : A \to C$ is injective. Then $g \circ h$ would be an injection of A into B, which we know cannot exist since A has greater cardinality than B. Hence it must be that no such injection h exists, which shows that A has greater cardinality than C as desired. □

(c) We define a sequence of sets recursively:

$$
\begin{aligned}
A_1 &= \mathbb{Z}_+\,, \\
A_n &= \mathcal{P}(A_{n-1}) \qquad \text{for } n > 1.
\end{aligned}
$$

We show that this meets the requirements.

Proof. First we show that each A_n is infinite by induction. Clearly $A_1 = \mathbb{Z}_+$ is infinite. Now assume that A_n is infinite for $n \in \mathbb{Z}_+$ so that there is an injection $f : \mathbb{Z}_+ \to A_n$ by Theorem 9.1. Then, by Lemma 9.7.1, $A_{n+1} = \mathcal{P}(A_n)$ has greater cardinality than A_n so that there is an injection $g : A_n \to A_{n+1}$. Then $g \circ f$ is an injection from $\mathbb{Z}_+$ to A_{n+1} so that A_{n+1} is infinite as well by Theorem 9.1. This completes the induction.

Finally, for any $n \in \mathbb{Z}_+$ we have that $n + 1 > 1$ so that $A_{n+1} = \mathcal{P}\left(A_{(n+1)-1}\right) = \mathcal{P}(A_n)$. Then clearly A_{n+1} has greater cardinality than A_n by Lemma 9.7.1. This shows the desired result. □

(d) Let $A = \bigcup_{n \in \mathbb{Z}_+} A_n$, which we claim has the required property.

Proof. Consider any $n \in \mathbb{Z}_+$. Clearly $A_n \subset A$ so that the identity function i_{A_n} is an injection of A_n into A. Now suppose for the moment that $g : A \to A_n$ is injective. Since clearly also $A_{n+1} \subset A$, it follows that $g \restriction A_{n+1}$ is then an injection of A_{n+1} into A_n. However this contradicts the proven fact that A_{n+1} has greater cardinality than A_n. Hence it has to be that no such injection g exists, which shows that A has greater cardinality than A_n. Since n was arbitrary, this shows the desired result. □

Exercise 9.8

Show that $\mathcal{P}(\mathbb{Z}_+)$ and $\mathbb{R}$ have the same cardinality. [Hint: You may use the fact that every real number has a decimal expansion, which is unique if expansions that end in an infinite string of 9's are forbidden.]

A famous conjecture of set theory, called the *continuum hypothesis*, asserts that there exists no set having cardinality greater than $\mathbb{Z}_+$ and lesser cardinality than $\mathbb{R}$. The *generalized continuum hypothesis* asserts that, given the infinite set A, there is no set having greater cardinality than A and lesser cardinality than $\mathcal{P}(A)$. Surprisingly enough, both of these assertions have been shown to be independent of the usual axioms of set theory. For a readable expository account, see [Sm].

Solution:

Lemma 9.8.1. *If A and B are sets with the same cardinality, then $\mathcal{P}(A)$ and $\mathcal{P}(B)$ have the same cardinality.*

Proof. Since A and B have the same cardinality there is a bijection $f : A \to B$. We define a bijection $g : \mathcal{P}(A) \to \mathcal{P}(B)$. So, for any $X \in \mathcal{P}(A)$, set $g(X) = f(X)$. Clearly $f(X) \subset B$, since B is the range of f, so that $g(X) = f(X) \in \mathcal{P}(B)$ and hence $\mathcal{P}(B)$ can be the range of g.

To show that g is injective, consider sets X and Y in $\mathcal{P}(A)$ so that $X, Y \subset A$. Also suppose that $X \neq Y$ so, without loss of generality, we can assume that there is an $x \in X$ where $x \notin Y$. Clearly $f(x) \in f(X)$ since $x \in X$. Were it the case that $f(x) \in f(Y)$ then there would be a $y \in Y$ such that $f(y) = f(x)$. But then we would have that $y = x$ since f is injective and hence $x = y \in Y$, which we know not to be the case. Hence $f(x) \notin f(Y)$ so that it has to be that $g(X) = f(X) \neq f(Y) = g(Y)$ since $f(x) \in f(X)$. Since X and Y were arbitrary this shows that g is injective.

To show that g is surjective consider any $Y \in \mathcal{P}(B)$ so that $Y \subset B$. Let $X = f^{-1}(Y)$, noting that f^{-1} is a bijection from B to A since f is bijective. Clearly $X \subset A$ since A is the range of f^{-1} so that $X \in \mathcal{P}(A)$. Now consider any $y \in f(X)$ so that there is an $x \in X$ where $f(x) = y$. Then, since $X = f^{-1}(Y)$, there is a $y' \in Y$ where $x = f^{-1}(y')$, and hence $y = f(x) = f(f^{-1}(y')) = y'$. Thus $y = y' \in Y$ so that $f(X) \subset Y$ since y was arbitrary. Now consider $y \in Y$ and let $x = f^{-1}(y)$ so that clearly $x = f^{-1}(y) \in f^{-1}(Y) = X$. Moreover, $f(x) = f(f^{-1}(y)) = y$ so that $y \in f(X)$. Thus $Y \subset f(X)$ as well since y was arbitrary. This shows that $g(X) = f(X) = Y$, from which we conclude that g is surjective since Y was arbitrary.

Hence $g : \mathcal{P}(A) \to \mathcal{P}(B)$ is a bijection so that $\mathcal{P}(A)$ and $\mathcal{P}(B)$ have the same cardinality by definition. $\square$

Main Problem.

Proof. We show this using the Cantor-Schroeder-Bernstein (CSB) Theorem, which was proven in Exercise 7.6b.

First, we construct an injective function f from $\mathbb{R}$ to $\mathcal{P}(\mathbb{Q})$. For any $x \in \mathbb{R}$ let $Q = \{q \in \mathbb{Q} \mid q < x\}$ so that clearly $Q \subset \mathbb{Q}$ and hence $Q \in \mathcal{P}(\mathbb{Q})$. Therefore setting $f(x) = Q$ means that f is a function from $\mathbb{R}$ to $\mathcal{P}(\mathbb{Q})$. To show that f is injective consider $x, y \in \mathbb{R}$ where $x \neq y$. Without loss of generality we can assume that $x < y$ so that there is a $q \in \mathbb{Q}$ where $x < q < y$ since the rationals are order-dense in the reals. Also set $Q = f(x)$ and $P = f(y)$. Since $q > x$ we have that $q \notin Q$. Analogously, since $q < y$ we have that $q \in P$. Thus it has to be that $f(x) = Q \neq P = f(y)$, which shows that f is injective since x and y were arbitrary.

Now, it was shown in Exercise 7.1 that $\mathbb{Q}$ is countably infinite and thus has the same cardinality as $\mathbb{Z}_+$. From Lemma 9.8.1 it then follows that $\mathcal{P}(\mathbb{Q})$ has the same cardinality as $\mathcal{P}(\mathbb{Z}_+)$ so that there is a bijection $g : \mathcal{P}(\mathbb{Q}) \to \mathcal{P}(\mathbb{Z}_+)$. Clearly then $g \circ f$ is an injection from $\mathbb{R}$ to $\mathcal{P}(\mathbb{Z}_+)$.

Now let $X = \{0, 1\}$, and we construct an injection $h : X^\omega \to \mathbb{R}$. For any sequence $\mathbf{x} = (x_1, x_2, \ldots) \in X^\omega$ set $h(\mathbf{x})$ to the decimal expansion $0.x_1x_2x_3\ldots$, where clearly each x_n is the digit 0 or 1. Clearly

$h(\mathbf{x})$ is a real number so that h is a function from X^ω to $\mathbb{R}$. It is easy to see that h is injective since different sequences will result in different decimal expansions. Since none of the expansions end in an infinite sequence of 9's, clearly the corresponding real numbers will be different.

Now, it was shown in Exercise 7.3 that $\mathcal{P}(\mathbb{Z}_+)$ and X^ω have the same cardinality so that there is a bijection $i : \mathcal{P}(\mathbb{Z}_+) \to X^\omega$. It then follows that $h \circ i$ is an injection of $\mathcal{P}(\mathbb{Z}_+)$ into $\mathbb{R}$. Since we have shown the existence of both injections, the result follows from the CSB Theorem. □

§10 Well-Ordered Sets

Exercise 10.1

Show that every well-ordered set has the least upper bound property.

Solution:

Proof. Suppose that A is a set with well-ordering $<$, and that B is some nonempty subset of A with upper bound $a \in A$. Let C then be the set of upper bounds of B, which is not empty since clearly $a \in C$. Then C is a nonempty subset of A and so has a smallest element c since A is well-ordered. Clearly then c is the least upper bound of B by definition. This shows that A has the least upper bound property since B was arbitrary. □

Exercise 10.2

(a) Show that in a well-ordered set, every element except the largest (if one exists) has an immediate successor.

(b) Find a set in which every element has an immediate successor that is not well-ordered.

Solution:

(a)

Proof. Suppose that A is well-ordered by $<$ and consider any $a \in A$ where a is not the largest element. It then follows that there is some $x \in A$ where $a < x$ since otherwise a would be the largest element of A. Let $X = \{y \in A \mid a < y\}$ so that clearly $X \subset A$ and $x \in X$. Thus X is a nonempty subset of A and so has a smallest element b since $<$ well-orders A. We claim that b is the immediate successor of a. To see this suppose that there is a $z \in A$ such that $a < z < b$, noting that clearly $a < b$ since $b \in X$. Then we would have that $z \in X$ but $z < b$ so that it is not true that $b \leq z$, which contradicts the definition of b as the smallest element of X. So it must be that no such z exists, which shows that b is indeed the immediate successor of a. □

(b) The most natural example of such a set is $\mathbb{Z}$. We show that this has the desired properties.

Proof. First, clearly $\mathbb{Z}$ is not well-ordered since, for example, the set of negative integers is a nonempty subset of $\mathbb{Z}$ but has no smallest element. Also, for any $n \in \mathbb{Z}$, clearly $n + 1$ is the immediate successor of n, which was shown back in Corollary 4.9.3. □

Exercise 10.3

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:

We claim that they do *not* have the same order type, which we show presently.

Proof. First, clearly $(1,1)$ is the smallest element of both ordered sets. For brevity let $A = \{1,2\} \times \mathbb{Z}_+$, $B = \mathbb{Z}_+ \times \{1,2\}$, and $<_A$ and $<_B$ be the corresponding dictionary orderings, with $<$ being the normal ordering of $\mathbb{Z}_+$.

So assume that they *do* have the same order type so that there is an order-preserving bijection $f : A \to B$. Consider $(2,1) \in A$, which is clearly not the smallest element since $(2,1) \neq (1,1)$. Let $(n,b) = f(2,1) \in B$, which cannot be the smallest element of B since f preserves order, so that $(n,b) \neq (1,1)$. Clearly $b \in \{1,2\}$ so that $b = 1$ or $b = 2$. In the former cases we must have that $n > 1$ so that $n - 1 \in \mathbb{Z}_+$. So set $y = (n-1, 2)$. In the latter case set $y = (n,1)$. It is easy to see, and trivial to formally show, that y is the immediate predecessor of (n,b) in either case.

Now let $x = f^{-1}(y)$, noting that f^{-1} is an order-preserving bijection from B to A since f is an order-preserving bijection. It then follows that $x <_A (2,1)$ since $f(x) = y <_B (n,b) = f(2,1)$. If $x = (m,a)$ then it has to be that $m < 2$ so that $m = 1$ (because $m \in \{1,2\}$) since there is no $a \in \mathbb{Z}_+$ where $a < 1$. Thus $x = (1,a)$ for some $a \in \mathbb{Z}_+$. We then have that $a + 1 \in \mathbb{Z}_+$ so that clearly $x = (1,a) <_A (1,a+1) <_A (2,1)$. From this we have, $y = f(1,a) <_B f(1,a+1) <_B f(2,1) = (n,b)$, which contradicts the fact that y is the immediate predecessor of (n,b). So it has to be that they do not have the same order type. $\square$

It is worth noting that, in the theory of ordinal numbers, $A = \{1,2\} \times \mathbb{Z}_+$ has order type $\omega + \omega = \omega \cdot 2$ whereas $B = \mathbb{Z}_+ \times \{1,2\}$ has simply order type ω. This also shows that A and B have different order types since distinct ordinal numbers always have different order types.

Exercise 10.4

(a) Let $\mathbb{Z}_-$ denote the set of negative integers in the usual order. Show that a simple 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 if A is simply ordered and every countable subset of A is well-ordered, then A is well-ordered.

Solution:

(a)

Proof. Let A be a set with simple order $\prec$.

($\Rightarrow$) Suppose that $\prec$ is not a well-ordering of A. Then there exists a nonempty subset B of A such that B has no smallest element. For any $b \in B$ define the set $X_b = \{x \in B \mid x \prec b\}$. Clearly $X_b \subset B$ and $X_b \neq \varnothing$ for any $b \in B$ since otherwise b would be the smallest element of B. Now let c be a choice function on the collection of nonempty subsets of B, which of course exists by the axiom of choice. Since B is nonempty there is a $b_0 \in B$. It then follows from the principle of recursive definition that there is a function $f : \mathbb{Z}_+ \to B$ such that

$$f(1) = b_0 \,,$$

$$f(n) = c(X_{f(n-1)}) \qquad \text{for } n > 1.$$

It then is easy to show that $f(n+1) \prec f(n)$ for all $n \in \mathbb{Z}_+$, i.e. that the sequence defined by f is decreasing. If we then simply define $g : \mathbb{Z}_- \to \mathbb{Z}_+$ by $g(n) = -n$ for $n \in \mathbb{Z}_-$, it is clear that $f \circ g$ is an order-preserving bijection from $\mathbb{Z}_-$ to some subset C of B. Clearly also $C \subset A$ since $B \subset A$. Hence the subset C has the same order type as $\mathbb{Z}_-$.

($\Leftarrow$) Now suppose that A has a subset B with the same order type as $\mathbb{Z}_-$. Clearly then B is nonempty and has no smallest element since $\mathbb{Z}_-$ does not. The existence of this B shows that A fails to be well-ordered. $\square$

(b)

Proof. Suppose that A is a set that is simply ordered by $\prec$ such that every countable subset is well-ordered by $\prec$. Consider any nonempty subset $B \subset A$. Suppose for a moment that the restricted $\prec$ does not well-order B. Then it follows from part (a) that B has a subset C with the same order type as $\mathbb{Z}_-$. However, clearly $C \subset A$ (since $B \subset A$) and C is countable (since $\mathbb{Z}_-$ is countable) and thus it should be well-ordered. As this is impossible since C has the same order-type as $\mathbb{Z}_-$ (which is clearly not well-ordered), it has to be that the restricted $\prec$ does in fact well-order B. Hence B has a $\prec$-smallest element, which shows that A is well-ordered since B was arbitrary. $\square$

Exercise 10.5

Show the well-ordering theorem implies the choice axiom.

Solution:

Proof. Suppose that $\mathcal{A}$ is a collection of nonempty sets. Then, by the well-ordering theorem there is a well-ordering $<$ of $\bigcup \mathcal{A}$. We construct a choice function $c : \mathcal{A} \to \bigcup \mathcal{A}$. Consider any set $A \in \mathcal{A}$. Since clearly A is then a nonempty subset of $\bigcup \mathcal{A}$, it follows that it has a unique smallest element a according to $<$ since $\bigcup \mathcal{A}$ is well-ordered by $<$. So simply set $c(A) = a$ so that clearly then $c(A) = a \in A$. This shows that c is in fact a choice function on $\mathcal{A}$. $\square$

Exercise 10.6

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 subset $\{x \mid \alpha < x\}$ 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:

Lemma 10.6.1. *If A is an uncountable set and $B \subset A$ is countable then $A - B$ is uncountable.*

Proof. If we let $C = A - B$, then clearly $A = C \cup B$. If C were countable then $A = C \cup B$ would be countable by Theorem 7.5 since B is also countable. Since we know that A is uncountable it therefore must be that $C = A - B$ is uncountable as well. $\square$

Main Problem.

It is assumed in the following that $<$ is the well-order on S_Ω.

(a)

Proof. Suppose to the contrary that S_Ω does have a largest element α. Then, for any $x \in S_\Omega$, we have that $x \leq \alpha$. Hence either $x \in \{y \in S_\Omega \mid y < \alpha\} = S_\alpha$ or $x = \alpha$. Therefore $S_\Omega = S_\alpha \cup \{\alpha\}$ since clearly both S_α and $\{\alpha\}$ are both subsets of S_Ω. Now since S_α is a section of S_Ω, it is countable. Since $\{\alpha\}$ is also clearly countable, it follows from Theorem 7.5 that their union $S_\alpha \cup \{\alpha\} = S_\Omega$ is countable. But this contradicts the fact that S_Ω is uncountable! Hence it has to be that S_Ω has no largest element as desired. $\square$

(b)

Proof. Consider any $\alpha \in S_\Omega$. Let $T_\alpha = \{x \in S_\Omega \mid \alpha < x\}$ so that we must show that T_α is uncountable. Let $\bar{S}_\alpha = S_\alpha \cup \{\alpha\}$ so that clearly we have that $\bar{S}_\alpha = \{x \in S_\Omega \mid x \leq \alpha\}$. It is then easy to show that $T_\alpha = S_\Omega - \bar{S}_\alpha$. Now, since S_α is a section of S_Ω, it is countable so that clearly $\bar{S}_\alpha = S_\alpha \cup \{\alpha\}$ is also countable by Theorem 7.5. Then, since S_Ω itself is uncountable, it follows that $T_\alpha = S_\Omega - \bar{S}_\alpha$ is also uncountable by Lemma 10.6.1. $\square$

(c)

Proof. First we show that X_0 is not bounded above. Assume the contrary so that $\alpha \in S_\Omega$ is an upper bound of X_0. It then follows that the set $T_\alpha = \{x \in S_\Omega \mid \alpha < x\}$ is such that every element of T_α has an immediate predecessor since otherwise there would be a $\beta \in T_\alpha$ where $\beta \in X_0$ so that α would not be an upper bound of X_0 since then $\alpha < \beta$.

Now, we know from part (a) that S_Ω has no largest element so that it follows from Exercise 10.2 that every element of S_Ω has an immediate successor. Since $T_\alpha \subset S_\Omega$ it follows that each element $x \in T_\alpha$ has an immediate successor y. Moreover we then have that $\alpha < x < y$ so that $y \in T_\alpha$ also. Hence every element of T_α has an immediate successor in T_α.

Now, we know that T_α is uncountable by part (b) so that it has a smallest element β since it is then a nonempty subset of the well-ordered S_Ω. We derive a contradiction by showing that T_α has the same order type as $\mathbb{Z}_+$ and is thus countable. We do this by defining an increasing bijection $f : \mathbb{Z}_+ \to T_\alpha$. First, set $f(1) = \beta$ and then set $f(n)$ to the immediate successor of $f(n-1)$ for $n > 1$, which was shown to exist above. Then the function f uniquely exists by the principle of recursive definition. Clearly we have that $f(n+1) > f(n)$ for all $n \in \mathbb{Z}_+$ since $f(n+1)$ is the immediate successor of $f(n)$. Hence f is increasing and therefore also injective.

To show that f is surjective suppose the contrary so that the set $T_\alpha - f(\mathbb{Z}_+)$ is nonempty. Since clearly this is a subset of the well-ordered S_Ω, it has a smallest element y. Now, we know that $f(1) = \beta$ so that $y \neq \beta$, and in fact $\beta < y$ since β is the smallest element of T_α. Since $y \in T_\alpha$ we know that it has an immediate predecessor x and that $\alpha < \beta \leq x$ so that $x \in T_\alpha$. However, it cannot be that $x \in T_\alpha - f(\mathbb{Z}_+)$ since $x < y$ and y is the smallest element of $T_\alpha - f(\mathbb{Z}_+)$. Thus $x \in f(\mathbb{Z}_+)$ so that there is an $n \in \mathbb{Z}_+$ where $f(n) = x$. But then $f(n+1) = y$ since y is the immediate successor of x. As this contradicts the fact that $y \notin f(\mathbb{Z}_+)$, it must be that f is in fact surjective!

Therefore we have shown that f is a bijection from $\mathbb{Z}_+$ to T_α so that T_α is countable. But we know from part (b) that T_α is uncountable. As mentioned above, this is a contradiction so that it must be that indeed X_0 is not bounded above. From this it immediately follows from the contrapositive of Theorem 10.3 that X_0 must be uncountable. $\square$

It is interesting to note that S_Ω corresponds to the ordinal number ω_1, which is the first uncountable ordinal, and the set X_0 of part (c) corresponds to the set of limit ordinals in ω_1. All of the curious properties deduced here for S_Ω apply to ω_1 too, assuming we allow the choice axiom.

Exercise 10.7

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 (The principle of transfinite induction). If J is a well-ordered set and J_0 is an inductive subset of J, then $J_0 = J$.

Solution:

Proof. Suppose that J_0 is an inductive subset of the well-ordered set J. Also suppose that $J_0 \neq J$. Since $J_0 \subset J$, it follows that there must be an $x \in J$ such that $x \notin J_0$. Thus the set $J - J_0$ is nonempty. Since clearly this is also a subset of J, it must have a smallest element α since J is well-ordered. Consider any $y \in S_\alpha$ so that $y < \alpha$. Then it cannot be that $y \in J - J_0$ since otherwise α would not be the smallest element of $J - J_0$. Since clearly $y \in J$ (since $S_\alpha \subset J$) it has to be that $y \in J_0$. Since y was arbitrary this shows that $S_\alpha \subset J_0$. It then follows that $\alpha \in J_0$ since J_0 is inductive. However, this contradicts the fact that $\alpha \in J - J_0$ so that our initial supposition that $J_0 \neq J$ must be incorrect. Hence $J_0 = J$ as desired. □

Exercise 10.8

(a) Let A_1 and A_2 be disjoint sets, well-ordered by $<_1$ and $<_2$, respectively. 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:

(a)

Proof. It is easy but tedious to show that $<$ is actually an order on $A_1 \cup A_2$, so we shall skip that proof and jump straight to the proof that it is a well-ordering.

So consider any nonempty subset A of $A_1 \cup A_2$.

Case: $A_1 \cap A \neq \varnothing$. Then clearly $A_1 \cap A$ is a nonempty subset of A_1 so that it has a smallest element a according to $<_1$ since it is a well-ordering. We then claim that a is the smallest element of A according to $<$. So consider any $x \in A$ so that clearly also $x \in A_1 \cup A_2$. Hence $x \in A_1$ or $x \in A_2$. If $x \in A_1$ then obviously $x \in A_1 \cap A$ so that $a \leq_1 x$ since a is the smallest element of $A_1 \cap A$. Then also $a \leq x$ by definition since a and x are both in A_1. On the other hand, if $x \in A_2$ then we again have that $a < x$ since $a \in A_1$ and $x \in A_2$. Therefore $a \leq x$ no matter what so that a is the smallest element of A since x was arbitrary.

Case: $A_1 \cap A = \varnothing$. Then it has to be that $A_2 \cap A \neq \varnothing$ since A is nonempty and $A = A_1 \cup A_2$. Thus $A_2 \cap A$ is a nonempty subset of A_2 so that it has a smallest element a by $<_2$ since it is a well-ordering. We claim that a is the smallest element of A. So consider any $x \in A$. It has to be that $x \in A_2$ since $A_1 \cap A$ is empty and $A = A_1 \cup A_2$. Therefore $x \in A_2 \cap A$ so that $a \leq_2 x$ since a is the smallest element of $A_2 \cap A$. Then, by definition, $a \leq x$ since both a and x are elements of A_2. This shows that a is the smallest element of A since x was arbitrary.

In either case we have shown that A has a smallest element so that $<$ is a well-ordering of $A_1 \cup A_2$ since A was arbitrary. □

Note that well-ordering a union of two well-ordered sets like this is analogous to the addition of two ordinal numbers. In particular if A_1 has order type α_1 and A_1 has order type α_2 where α_1 and α_2 are an ordinal numbers, then $A_1 \cup A_2$ with the above well-ordering has order type $\alpha_1 + \alpha_2$.

(b) Suppose that J is well-ordered by $<_J$ and $\{A_\alpha\}_{\alpha \in J}$ is a collection of well-ordered sets where A_α is well-ordered by $<_\alpha$ for each $\alpha \in J$. Now define an order $<$ on $A = \bigcup_{\alpha \in J} A_\alpha$ as follows. For any x and y in A there are clearly α and β in J where $x \in A_\alpha$ and $y \in A_\beta$, noting that α and β are unique since the collection is mutually disjoint. So set $x < y$ if and only if either $\alpha = \beta$ and $x <_\alpha y$, or else $\alpha <_J \beta$, noting that these are clearly mutually exclusive. We then claim that $<$ is a well-ordering of A.

Proof. Let B be any nonempty subset of A and I be the set of $\alpha \in J$ such that there is an $x \in B$ where $x \in A_\alpha$. Now, since B is nonempty, there is a $z \in B$. Since $B \subset A = \bigcup_{\alpha \in J} A_\alpha$, there is an $\gamma \in J$ where $z \in A_\gamma$. Then clearly $\gamma \in I$ so that I is a nonempty subset of J. Then I has a smallest element α since it is well-ordered by $<_J$. By the definition of I there is a $w \in B$ where $w \in A_\alpha$. Then clearly $w \in A_\alpha \cap B$ so that it is a nonempty subset of A_α. It then follows that $A_\alpha \cap B$ has a smallest element a according to $<_\alpha$ since it is a well-ordering on A_α. We claim that a is the smallest element of B by $<$.

So consider any $x \in B$ so that there is a $\beta \in J$ where $x \in A_\beta$ since $B \subset A$.

Case: $\beta = \alpha$. Then both a and x are in $A_\alpha \cap B = A_\beta \cap B$ so that $a \leq_\alpha b$ since a is the smallest element of $A_\alpha \cap B$. It then follows from the definition of $<$ that $a \leq x$.

Case: $\beta \neq \alpha$. Clearly then $\beta \in I$ so that $\alpha \leq_J \beta$ since α is the smallest element of J. Since we know that $\beta \neq \alpha$ it must be that $\alpha <_J \beta$. From this it follows that $a < x$ by definition.

Hence in either case it is true that $a \leq x$, which shows that a is the smallest element of B. Since B was an arbitrary nonempty subset of A, this shows that A is well-ordered by $<$. $\square$

Exercise 10.9

Consider the subset A of $(\mathbb{Z}_+)^\omega$ consisting of all infinite sequences of positive integers $\mathbf{x} = (x_1, x_2, \ldots)$ that end in an infinite string of 1's. Give A the following order: $\mathbf{x} < \mathbf{y}$ if $x_n < y_n$ and $x_i = y_i$ for $i > n$. We call this the "antidictionary order" on A.

(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:

(a)

Proof. Consider any positive integer n. Define a sequence $\mathbf{a} = (a_1, a_2, \ldots)$ in A by

$$a_i = \begin{cases} 2 & i = n+1 \\ 1 & i \neq n+1 \,. \end{cases}$$

We claim that the section $S_{\mathbf{a}}$ has the same order type as $(\mathbb{Z}_+)^n$. To show this we construct an order-preserving mapping $f : S_{\mathbf{a}} \to (\mathbb{Z}_+)^n$. So consider any sequence $\mathbf{x} = (x_1, x_2, \ldots)$ in $S_{\mathbf{a}}$ so that $\mathbf{x} < \mathbf{a}$. Now define a finite sequence where $y_i = x_{n-i+1}$ for any $i \in \{1, \ldots, n\}$, and set $f(\mathbf{x}) = \mathbf{y} = (y_1, \ldots, y_n)$. Clearly $f(\mathbf{x}) \in (\mathbb{Z}_+)^n$ since $\mathbf{x} \in A \subset (\mathbb{Z}_+)^\omega$.

Here we must digress for a moment and show that, for all $\mathbf{x} = (x_1, x_2, \ldots) \in A$, $\mathbf{x} \in S_{\mathbf{a}}$ if and only if $x_i = 1$ for all $i > n$.

($\Rightarrow$) We show the contrapositive. So suppose that there is an $i > n$ where $x_i \neq 1$. Moreover let i be the greatest such index, which must exist since $\mathbf{x}$ must end in an infinite string of 1's. Clearly then the fact that $x_i \in \mathbb{Z}_+$ and $x_i \neq 1$ means that $x_i > 1$. Now, if $i > n+1$ then we have that $x_i > 1 = a_i$ and $x_j = 1 = a_j$ for all $j > i$ so that clearly $\mathbf{x} > \mathbf{a}$. If $i = n+1$ and $i > 2$ clearly $x_i > 2 = a_{n+1} = a_i$ and $x_j = 1 = a_j$ for all $j > i$ so that again $\mathbf{x} > \mathbf{a}$. Lastly suppose that $i = n+1$ but that $x_i = 2 = a_{n+1} = a_i$. If $x_j = 1 = a_j$ for all $j < i = n+1$ then clearly $\mathbf{x} = \mathbf{a}$. On the other hand if there is a $1 \leq j < n+1 = i$ where $x_j \neq 1$ then let j be the greatest such index. Then we clearly have $x_j > 1 = a_j$ while $x_k = a_k$ for all $k > j$ so that $\mathbf{x} > \mathbf{a}$. Therefore in every one of these exhaustive cases we have that $\mathbf{x} \geq \mathbf{a}$ so that $\mathbf{a} \notin S_{\mathbf{a}}$.

($\Leftarrow$) Now suppose that $x_i = 1$ for every $i > n$. Then we have that $x_{n+1} = 1 < 2 = a_{n+1}$ while $x_j = 1 = a_j$ for all $j > n+1 > n$ so that $\mathbf{x} < \mathbf{a}$ and hence $\mathbf{x} \in S_{\mathbf{a}}$.

Now, returning to the main proof, we first show that f as defined above preserves order. To this end let $\prec$ denote the dictionary order on $(\mathbb{Z}_+)^n$. Now consider any $\mathbf{x} = (x_1, x_2, \ldots)$ and $\mathbf{x}' = (x_1', x_2', \ldots)$ in $S_{\mathbf{a}}$ where $\mathbf{x} < \mathbf{x}'$. Also let $\mathbf{y} = f(\mathbf{x})$ and $\mathbf{y}' = f(\mathbf{x}')$. Then there is an $m \in \mathbb{Z}_+$ where $x_m < x_m'$ and $x_i = x_i'$ for all $i > m$. We also have by what was shown above that $x_i = 1 x_i'$ for all $i > n$ since $\mathbf{x}, \mathbf{x}' \in S_{\mathbf{a}}$. So it has to be that $m \leq n$. It then follows from the fact that $1 \leq m \leq n$ that $1 \leq n - m + 1 \leq n$ as well. Thus we have

$$y_{n-m+1} = x_{n-(n-m+1)+1} = x_m < x_m' = x_{n-(n-m+1)+1}' = y_{n-m+1}' .$$

For any $1 \leq j < n - m + 1$ we have that $n - j + 1 > m$ so that

$$y_j = x_{n-j+1} = x_{n-j+1}' = y_j' .$$

Thus by definition we have that $f(\mathbf{x}) = \mathbf{y} \prec \mathbf{y}' = f(\mathbf{x}')$, which shows that f preserves order since $\mathbf{x}$ and $\mathbf{x}'$ were arbitrary. Note that this also clearly shows that f is injective.

To show that f is also surjective, consider any $\mathbf{y} = (y_1, \ldots, y_n) \in (\mathbb{Z}_+)^n$. Now define a sequence

$$x_i = \begin{cases} y_{n-i+1} & 1 \leq i \leq n \\ 1 & i > n \end{cases}$$

so that clearly $\mathbf{x} = (x_1, x_2, \ldots) \in S_{\mathbf{a}}$ by what was shown above. Now let $\mathbf{y}' = (y_1', \ldots y_n') = f(\mathbf{x})$. Consider any $1 \leq i \leq n$ and let $j = n - i + 1$ so that also $n - j + 1 = i$, noting also that $1 \leq j \leq n$. Then we have

$$y_i = y_{n-j+1} = x_j = x_{n-i+1} = y_i'$$

by the definition of f. Since i was arbitrary this shows that $f(\mathbf{x}) = \mathbf{y}' = \mathbf{y}$, which shows that f is surjective since $\mathbf{y}$ was arbitrary.

The existence of f therefore shows that $S_{\mathbf{a}}$ and $(\mathbb{Z}_+)^n$ have the same order type. $\square$

(b)

Proof. Consider any nonempty subset B of A. Clearly the sequence $(1, 1, \ldots)$ is the smallest element of A and hence if it is in B then it is also the smallest element of B. So suppose that $(1, 1, \ldots) \notin B$ so that, for every $\mathbf{x} \in B$ there is a unique greatest $n_{\mathbf{x}} \in \mathbb{Z}_+$ where $x_{n_{\mathbf{x}}} > 1$ but $x_i = 1$ for all $i > n_{\mathbf{x}}$. So let $I = \{n_{\mathbf{x}} \mid \mathbf{x} \in B\}$, noting that $B \neq \varnothing$ implies that $I \neq \varnothing$ as well. Thus I is a nonempty subset of $\mathbb{Z}_+$ and hence has a smallest element n. If we then let B_n be the set of sequences $\mathbf{x} \in B$

where $x_n > 1$ but $x_i = 1$ for all $i > n$, then the fact that $n \in I$ clearly implies that $B_n \neq \varnothing$. Also, if we define the sequence

$$a_i = \begin{cases} 2 & i = n+1 \\ 1 & i \neq n+1 \end{cases}$$

as in part (a) then it follows from what was shown there that $B_n \subset S_{\mathbf{a}}$. Moreover it was shown that $S_{\mathbf{a}}$ has the same order type as the dictionary order of $(\mathbb{Z}_+)^n$, which we know to be a well-ordering. Hence $S_{\mathbf{a}}$ must also be a well-ordering so that B_n has a smallest element $\mathbf{b} = (b_1, b_2, \ldots)$ since it is a nonempty subset of $S_{\mathbf{a}}$. We claim that $\mathbf{b}$ is in fact the smallest element of all of B.

So consider any $\mathbf{x} \in B$ so that $n_{\mathbf{x}} \in I$. It then follows that $n \leq n_{\mathbf{x}}$ since it is the smallest element of I. If $n = n_{\mathbf{x}}$ then we have that $\mathbf{x} \in B_n$ so that $\mathbf{b} \leq \mathbf{x}$ since it is the smallest element of B_n. If $n < n_{\mathbf{x}}$ then we have that $b_{n_{\mathbf{x}}} = 1 < x_{n_{\mathbf{x}}}$ but $b_i = 1 = x_i$ for every $i > n_{\mathbf{x}} > n$. This shows that $\mathbf{b} < \mathbf{x}$. Thus in all cases $\mathbf{b} \leq \mathbf{x}$, which shows that $\mathbf{b}$ is the smallest element of B since $\mathbf{x}$ was arbitrary. Since B was arbitrary, this shows that A is well-ordered as desired. $\square$

Note that, in the theory of ordinal numbers, the set $(\mathbb{Z}_+)^n$ (and therefore the corresponding section of A) has order type ω^n. It would seem then that the set A has order type ω^ω.

Exercise 10.10

Theorem. Let J and C be well-ordered sets; assume that there is no surjective function mapping a section of J onto C. Then there exists a unique function $h : J \to C$ satisfying the equation

$$(*) \qquad\qquad h(x) = \text{smallest}[C - h(S_x)]$$

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 $(*)$ 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 $(*)$, show that there exists a function $k : S_\alpha \cup \{\alpha\} \to C$ satisfying $(*)$.

(c) If $K \subset J$ and for all $\alpha \in K$ there exists a function $h_\alpha : S_\alpha \to C$ satisfying $(*)$, show that there exists a function

$$k : \bigcup_{\alpha \in K} S_\alpha \to C$$

satisfying $(*)$.

(d) Show by transfinite induction that for every $\beta \in J$, there exists a function $h_\beta : S_\beta \to C$ satisfying $(*)$. [Hint: If β has an immediate predecessor α, then $S_\beta = S_\alpha \cup \{\alpha\}$. If not, S_β is the union of all S_α with $\alpha < \beta$.]

(e) Prove the theorem.

Solution:

The following lemma is proof by transfinite induction, which is more straightforward than having to frame everything in terms of inductive sets. Henceforth we use this whenever transfinite induction is required.

Lemma 10.10.1. *(Proof by transfinite induction) Suppose that J is a well-ordered set and $P(x)$ is a proposition with parameter x. Suppose also that if $P(x)$ is true for all $x \in S_\alpha$ (where S_α is a section of J), then $P(\alpha)$ is also true. Then $P(\beta)$ is true for every $\beta \in J$.*

Proof. Let $J_0 = \{x \in J \mid P(x)\}$. We show that J_0 is inductive. So consider any $\alpha \in J$ and suppose that $S_\alpha \subset J_0$. Then, for any $x \in S_\alpha$ we have that $x \in J_0$ so that $P(x)$. It then follows that $P(\alpha)$ is also true since x was arbitrary, and so $\alpha \in J_0$. Since $\alpha \in J$ was arbitrary, this shows that J_0 is inductive. It then follows from Exercise 10.7 that $J_0 = J$. So consider any $\beta \in J$ so that also $\beta \in J_0$ and hence $P(\beta)$ is true. Since β was arbitrary, this shows the desired result. $\square$

Main Problem.

(a)

Proof. First suppose that the domains of h and k are sets H and K where each is either a section of J or J itself. Since this is the case, we can assume without loss of generality that $H \subset K$ and so H is exactly the domain common to both h and k. Now suppose that the hypothesis we are trying to prove is *not* true so that there is an x in both domains (i.e. $x \in H$) where $h(x) \neq k(x)$. We can also assume that x is the smallest such element since $H \subset J$ and J is well-ordered. It then clearly follows that $S_x \subset H$ is a section of J and that $h(y) = k(y)$ for all $y \in S_x$. From this we clearly have that $h(S_x) = k(S_x)$. But then we have

$$h(x) = \text{smallest}[C - h(S_x)] = \text{smallest}[C - k(S_x)] = k(x)$$

since both h and k satisfy $(*)$ and x is in the domain of both. This contradicts the supposition that $h(x) \neq k(x)$ so that it must be that no such x exists and hence h and k are the same in their common domain as desired. $\square$

(b)

Proof. Suppose that $h : S_\alpha \to C$ is such a function satisfying $(*)$. Now let $\bar{S}_\alpha = S_\alpha \cup \{\alpha\}$ and we define $k : \bar{S}_\alpha \to C$ as follows. For any $x \in \bar{S}_\alpha$ set

$$k(x) = \begin{cases} h(x) & x \in S_\alpha \\ \text{smallest}[C - h(S_\alpha)] & x = \alpha\,. \end{cases}$$

We note that clearly S_α and $\{\alpha\}$ are disjoint so that this is unambiguous. We also note that h is not surjective onto C since S_α is a section of J, and hence $C - h(S_\alpha) \neq \varnothing$ and so has a smallest element since C is well-ordered.

Now we show that k satisfies $(*)$. First, clearly $h(S_x) = k(S_x)$ for any $x \leq \alpha$ since $k(y) = h(y)$ by definition for any $y \in S_x \subset S_\alpha$. Now consider any $x \in \bar{S}_\alpha$. If $x = \alpha$ then by definition we have

$$k(x) = \text{smallest}[C - h(S_\alpha)] = \text{smallest}[C - k(S_\alpha)] = \text{smallest}[C - k(S_x)]\,.$$

On the other hand, if $x \in S_\alpha$ then $x < \alpha$ so that

$$k(x) = h(x) = \text{smallest}[C - h(S_x)] = \text{smallest}[C - k(S_x)]$$

since h satisfies $(*)$. Therefore, since x was arbitrary, this shows that k also satisfies $(*)$. $\square$

(c)

Proof. Let

$$k = \bigcup_{\alpha \in K} h_\alpha \,,$$

which we claim is the function we seek.

First we show that k is actually a function from $\bigcup_{\alpha \in K} S_\alpha$ to C. So consider any x in the domain of k. Suppose that (x,a) and (x,b) are both in k so that there are α and β in K where $(x,a) \in h_\alpha$ and $(x,b) \in h_\beta$. Since h_α and h_β both satisfy $(*)$, it follows from part (a) that $a = h_\alpha(x) = h_\beta(x) = b$ since clearly x is in the domain of both. This shows that k is indeed a function since (x,a) and (x,b) were arbitrary. Also clearly the domain of k is $\bigcup_{\alpha \in K} S_\alpha$ since, for any $x \in \bigcup_{\alpha \in K} S_\alpha$, we have that there is an $\alpha \in K$ where $x \in S_\alpha$. Hence x is in the domain of h_α and so in the domain of k. In the other direction, clearly if x is in the domain of k then it is in the domain of h_α for some $\alpha \in K$. Since this domain is S_α, clearly $x \in \bigcup_{\alpha \in K} S_\alpha$. Lastly, obviously the range of k can be C since this is the range of every h_α.

Now we show that k satisfies $(*)$. So consider any $x \in \bigcup_{\alpha \in K} S_\alpha$ so that $x \in S_\alpha$ for some $\alpha \in K$. Clearly we have that $k(y) = h_\alpha(y)$ for every $y \in S_\alpha$ since $h_\alpha \subset k$. It then immediately follows that $k(x) = h(x)$ and $k(S_x) = h_\alpha(S_x)$ since $S_x \subset S_\alpha$. Then, since h_α satisfies $(*)$, we have

$$k(x) = h_\alpha(x) = \text{smallest}[C - h_\alpha(S_x)] = \text{smallest}[C - k(S_x)] \,.$$

Since x was arbitrary, this shows that k satisfies $(*)$ as desired. □

(d)

Proof. Consider any $\beta \in J$ and suppose that, for every $x \in S_\beta$, there is a function $h_x : S_x \to C$ satisfying $(*)$. Now, if β has an immediate predecessor α then we claim that $S_\beta = S_\alpha \cup \{\alpha\}$. First if $x \in S_\beta$ then $x < \beta$ so that $x \leq \alpha$ since α is the immediate predecessor of β. If $x < \alpha$ then $x \in S_\alpha$ and if $x = \alpha$ then $x \in \{\alpha\}$. Hence in either case we have that $x \in S_\alpha \cup \{\alpha\}$. Now suppose that $x \in S_\alpha \cup \{\alpha\}$. If $x \in S_\alpha$ then $x < \alpha < \beta$ so that $s \in S_\beta$. On the other hand if $x \in \{\alpha\}$ then $x = \alpha < \beta$ so that again $x \in S_\beta$. Thus we have shown that $S_\beta \subset S_\alpha \cup \{\alpha\}$ and $S_\alpha \cup \{\alpha\} \subset S_\beta$ so that $S_\beta = S_\alpha \cup \{\alpha\}$. Since $\alpha \in S_\beta$ it follows that there is an $h_\alpha : S_\alpha \to C$ that satisfies $(*)$. Then, by part (b), we have that there is an $h_\beta : S_\beta = S_\alpha \cup \{\alpha\} \to C$ that also satisfies $(*)$.

If β does not have an immediate predecessor then we claim that $S_\beta = \bigcup_{\gamma < \beta} S_\gamma$. So consider any $x \in S_\beta$ so that $x < \beta$. Since x cannot be the immediate predecessor of β, there must be an α where $x < \alpha < \beta$. Then $x \in S_\alpha$ so that, since $\alpha < \beta$, clearly $x \in \bigcup_{\gamma < \beta} S_\gamma$. Now suppose that $x \in \bigcup_{\gamma < \beta} S_\gamma$ so that there is an $\alpha < \beta$ where $x \in S_\alpha$. Then clearly $x < \alpha < \beta$ so that also $x \in S_\beta$. Thus we have shown that $S_\beta \subset \bigcup_{\gamma < \beta} S_\gamma$ and $\bigcup_{\gamma < \beta} S_\gamma \subset S_\beta$ so that $S_\beta = \bigcup_{\gamma < \beta} S_\gamma$. Now, clearly S_β is a subset of J where there is an $h_x : S_x \to C$ satisfying $(*)$ for every $x \in S_\beta$. Then it follows from what was shown in part (c) that there is a function h_β from $\bigcup_{\gamma < S_\beta} S_\gamma = \bigcup_{\gamma < \beta} S_\gamma = S_\beta$ to C that satisfies $(*)$.

Therefore, in either case, we have shown that there is an $h_\beta : S_\beta \to C$ that satisfies $(*)$. The desired result then follows by transfinite induction. □

(e)

Proof. First suppose that J has no largest element. Then we claim that $J = \bigcup_{\alpha \in J} S_\alpha$. For any $x \in J$ there must be a $y \in J$ where $x < y$ since x cannot be the greatest element of J. Hence $x \in S_y$ so that also clearly $\bigcup_{\alpha \in J} S_\alpha$. Then, for any $x \in \bigcup_{\alpha \in J} S_\alpha$, there is an $\alpha \in J$ where $x \in S_\alpha$. Clearly $S_\alpha \subset J$ so that $x \in J$ also. Hence $J \subset \bigcup_{\alpha \in J} S_\alpha$ and $\bigcup_{\alpha \in J} S_\alpha \subset J$ so that $J = \bigcup_{\alpha \in J} S_\alpha$. Since we know from part (d) that there is an $h_\alpha : S_\alpha \to C$ that satisfies $(*)$ for every $\alpha \in J$, it follows from part (c) that there is a function h from $\bigcup_{\alpha \in J} S_\alpha = J$ to C that satisfies $(*)$.

If J does have a largest element β then clearly $J = S_\beta \cup \{\beta\}$. Since we know that there is an $h_\beta : S_\beta \to C$ that satisfies $(*)$ by part (d), it follows from part (b) that there is a function h from $S_\beta \cup \{\beta\} = J$ to C that satisfies $(*)$. Hence the desired function h exists in both cases. Part (a) also clearly shows that this function is unique. □

Exercise 10.11

Let A and B be two sets. Using the well-ordering theorem, prove that either they have the same cardinality, or one has cardinality greater than the other. [Hint: If there is no surjection $f : A \to B$, apply the preceding exercise.]

Solution:

Lemma 10.11.1. *For well-ordered sets $A \neq \varnothing$ and B there is an injection from A to B if and only if there is a surjection from B to A.*

Proof. ($\Rightarrow$) Suppose that there is an injection $f : A \to B$ and $A \neq \varnothing$. Then there is an $a \in A$. We then construct a surjection $g : B \to A$ as follows. For any $y \in B$ if $b \in f(A)$ then there is a unique $x \in A$ where $y = f(x)$. It is unique since, if x and x' are in A where $f(x) = y = f(x')$, then $x = x'$ since f is injective. So in this case set $g(y) = x$. If $b \notin f(A)$, then set $g(y) = a$. Clearly g is a function from B to A. To show that g is surjective, consider any $x \in A$ and let $y = f(x)$, which is clearly an element of B. Then, since obviously $y \in f(A)$ and x is the unique $x \in A$ such that $y = f(x)$, we have that $g(y) = x$ by definition. This shows that g is surjective since x was arbitrary.

($\Leftarrow$) Now suppose that $g : B \to A$ is surjective. We then construct an injection $f : A \to B$ as follows. For any $x \in A$ we have that the set $B_x = \{y \in B \mid g(y) = x\}$ is nonempty since g is surjective. Hence B_x has a unique smallest element y since it is a nonempty subset of B and B is well-ordered. So simply set $f(x) = y$. Clearly f is a function from A to B. To show that f is injective, consider $x, x' \in A$ where $x \neq x'$. Then clearly the sets B_x and $B_{x'}$ have to be disjoint for otherwise there would be a $y \in B$ where $g(y) = x$ and $g(y) = x'$, which is impossible if $x \neq x'$ since g is a function. Hence, since $f(x)$ and $f(x')$ are defined to be the smallest elements of B_x and $B_{x'}$, respectively, we have $f(x) \neq f(x')$. This shows that f is injective since x and x' were arbitrary. □

Main Problem.

Proof. First suppose that A and B are each well-ordered, which follows from the well-ordering theorem. Also suppose that A and B do *not* have the same cardinality so that it suffices to show that either B has greater cardinality than A or vice versa. If $A = \varnothing$ then it cannot be that $B = \varnothing$ as well since then they would have the same cardinality ($\varnothing$ would be a trivial bijection between them). Hence $B \neq \varnothing$ so that clearly B has greater cardinality than A. Thus in what follows assume that $A \neq \varnothing$.

Suppose that there is an injection from A to B. Then there cannot be an injection from B to A since, if there were, then A and B would have the same cardinality by the Cantor-Schroeder-Bernstein Theorem (shown in Exercise 7.6b). Thus B has greater cardinality than A by definition.

On the other hand, if there is no injection from A to B then there is no surjection from B to A by Lemma 10.11.1 since they are both well-ordered and $A \neq \varnothing$. It then clearly follows that no section of B can be a surjection onto A since then any extension of such a function to all of B would also be a surjection onto A. From this we have by Exercise 10.10 that there is a unique function $h : B \to A$ with the property that

$$h(x) = \text{smallest}[A - h(S_x)]\,,$$

where of course S_x is the section of B by x.

We claim that h is injective. So consider any y and y' in B where $y \neq y'$. Without loss of generality we can assume that $y < y'$ (by the well-ordering on B). It then follows that $y \in S_{y'}$ so that clearly $h(y) \in h(S_{y'})$. However, we have that $h(y')$ is the smallest element of $A - h(S_{y'})$ so that obviously $h(y') \notin h(S_{y'})$. Hence it must be that $h(y) \neq h(y')$, which shows that h is injective since y and y' were arbitrary.

Therefore there is an injection from B to A but none from A to B so that A has greater cardinality than B by definition. This shows the desired result since these cases are exhaustive. □

§11 The Maximum Principle

Exercise 11.1

If a and b are real numbers, define $a \prec b$ if $b - a$ is positive and rational. Show this is a strict partial order on $\mathbb{R}$. What are the maximal simply ordered subsets?

Solution:

Lemma 11.1.1. *If B is a maximal simply ordered subset of a nonempty partially ordered set A, then B is nonempty.*

Proof. Since A is nonempty, there is an $a \in A$. Clearly $\varnothing$ is vacuously simply ordered. However, it cannot be maximal since clearly the set $\{a\}$ properly contains $\varnothing$ as a subset but is also clearly vacuously simply ordered by $\prec$. Hence, since B *is* maximal it must be that $B \neq \varnothing$ as desired. □

Main Problem.

First we show that $\prec$ is a strict partial order.

Proof. First consider any $a \in \mathbb{R}$ so that $a - a = 0$, which is not positive and hence it is not true that $a \prec a$. Therefore $\prec$ is nonreflexive. Now consider $a, b, c \in \mathbb{R}$ where $a \prec b$ and $b \prec c$. Then we have that $x = b - a$ and $y = c - b$ are positive and rational. It then clearly follows that

$$c - a = (c - b) + (b - a) = y + x$$

is also rational and positive since both x and y are. Thus $a \prec c$, which shows that $\prec$ is transitive. Since $\prec$ was shown to be nonreflexive and transitive, this shows that it is a strict partial order as desired. □

For any element $x \in \mathbb{R}$, define the set $A_x = \{y \in \mathbb{R} \mid x - y \in \mathbb{Q}\}$. We then claim that the collection $\mathcal{A} = \{A_x\}_{x \in \mathbb{R}}$ is exactly the set of all maximal simply ordered subsets.

Proof. Suppose that $\mathcal{B}$ is the set of maximally simply ordered subsets of $\mathbb{R}$. Then we show that $\mathcal{A} = \mathcal{B}$.

To show that $\mathcal{A} \subset \mathcal{B}$ consider any $X \in \mathcal{A}$ so that $X = A_x$ for some $x \in \mathbb{R}$. Now consider any distinct y and z in $X = A_x$ so that by definition $x - y$ and $x - z$ are both rational so that $z - x = -(x - z)$ is also rational. Then clearly $z - y = (z - x) + (x - y)$ is rational as is $y - z = -(z - y)$. Since y and z are distinct, we have that $z - y$ and $y - z$ are nonzero and that either $y < z$ or $z < y$. In the former case we have that $z - y$ is a positive rational number and in the latter $y - z$ is. Thus either $y \prec z$ or $z \prec y$, which shows that $X = A_x$ is simply ordered since y and z were arbitrary. Now consider any

$y \in A_x$ and $z \notin A_x$ so that $x - y$ is rational but $x - z$ is irrational so that $z - x = -(x - z)$ is also irrational. Since a rational added to an irrational is also irrational (which is trivially easy to prove), it follows that $z - y = (z - x) + (x - y)$ is irrational as is $y - z = -(z - y)$. Hence it cannot be that either $y \prec z$ or $z \prec y$. Since $y \in X$ and $z \notin X$ were arbitrary, this show that X is a maximal simply ordered set so that $X \in \mathcal{B}$. This shows that $\mathcal{A} \subset \mathcal{B}$ since X was arbitrary.

Now suppose that $X \in \mathcal{B}$ so that X is a maximal simply ordered set. It follows from Lemma 11.1.1 that X is nonempty so that there is an $x \in X$, and we claim that in fact $X = A_x$. So consider any $y \in X$. Clearly if $y = x$ then $x - y = x - x = 0 \in \mathbb{Q}$ so that $y \in A_x$. If $y \neq x$ than either $x \prec y$ or $y \prec x$ since X is simply ordered by $\prec$. In the former case we have that $y - x$ is positive and rational so that $x - y = -(y - x)$ is negative and rational, and hence $y \in A_x$. In the latter case we have that $x - y$ is positive and rational so that clearly again $y \in A_x$. Since y was arbitrary this shows that $X \subset A_x$. Now consider any $y \in A_x$ so that $x - y \in \mathbb{Q}$. If $y = x$ then clearly $y \in X$. If $y \neq x$ then either $y - x$ or $x - y$ is positive, and also clearly rational since $x - y$ is rational. Hence either $x \prec y$ or $y \prec x$. It then follows from the fact that X is *maximally* simply ordered that y must be in X since otherwise y would not be comparable with x. Since again y was arbitrary this shows that $A_x \subset X$. Hence $X = A_x$ so that clearly $X \in \{A_x\}_{x \in \mathbb{R}} = \mathcal{A}$. Since X was arbitrary this shows that $\mathcal{B} \subset \mathcal{A}$.

Therefore we have shown that $\mathcal{A} = \mathcal{B}$, which shows that $\mathcal{A}$ is exactly the complete set of maximally simply ordered subsets. □

As an example of a particular maximally well-ordered set we have $\mathbb{Q} = A_0$ itself.

Exercise 11.2

(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:

(a)

Proof. We show that $\preceq$ satisfies the three partial order axioms:

(i) Consider any $a \in A$. Since obviously $a = a$ we have by definition that $a \preceq a$.

(ii) Suppose that $a \preceq b$ and $b \preceq a$. Then either $a \prec b$ or $a = b$, and either $b \prec a$ or $b = a$. So suppose that $a \neq b$ so that it must be that $a \prec b$ and $b \prec a$. Since $\prec$ is a strict partial order, it is transitive so that $a \prec a$ since $a \prec b$ and $b \prec a$. But this contradicts the nonreflexivity of $\prec$. Hence it must be that $a = b$ as desired.

(iii) Suppose that $a \preceq b$ and $b \preceq c$. Hence either $a \prec b$ or $a = b$, and either $b \prec c$ or $b = c$.

Case: $a \prec b$. If $b \prec c$ then clearly $a \prec c$ since $\prec$ is transitive (since it is a strict partial order). If $b = c$ then we have that $a \prec b = c$.

Case: $a = b$. If $b \prec c$ then we have that $a = b \prec c$. If $b = c$ then we have that $a = b = c$.

Hence in all cases and sub-cases we have that $a \prec c$ or $a = c$, and thus $a \preceq c$ by definition. $\square$

(b)

Proof. We show that S satisfies the two strict partial order axioms:

Nonreflexivity. Consider any $a \in A$. Since $a = a$ it follows that it is not true that $a \neq a$ and hence not true that aSa. Thus S is nonreflexive since a was arbitrary.

Transitivity. Suppose that aSb and bSc. Hence by definition aPb and $a \neq b$, and bPc and $b \neq c$. Then, by the transitivity property of the partial order axioms, which is property (iii), we have that aPc. Suppose for a moment that $a = c$. Then we would have aPb and bPa (since bPc and $c = a$). Then by partial order axiom (ii) we have that $a = b$, which contradicts the fact that $a \neq b$. So it must be that $a \neq c$. Thus aPc and $a \neq c$ so that aSc, which shows that S is transitive. $\square$

Exercise 11.3

Let A be a set with 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 a *comparable* with x:

$$B = \{y \mid y \in A \text{ and either } x \prec y \text{ or } y \prec x\} .$$

But this will not always work. In which of Examples 1 and 2 will this procedure succeed and in which will it not?

Solution:

First, it seems that, as defined above, B does not actually contain x itself! This is because it is not true that $x \prec x$ by the nonreflexivity of the partial order $\prec$. We assume that this was an oversight, which is easily remedied by defining

$$B' = \{y \in A \mid \text{either } x \prec y \text{ or } y \prec x\}$$

and $B = B' \cup \{x\}$.

For Example 1, a circular region in $\mathbb{R}^2$ is clearly

$$C_{\mathbf{x}_0,r} = \left\{\mathbf{x} \in \mathbb{R}^2 \mid |\mathbf{x} - \mathbf{x}_0| < r\right\} ,$$

where the point $\mathbf{x}_0 \in \mathbb{R}^2$ is the center of the circle, $r \in \mathbb{R}_+$ is the radius, and $|(x, y)| = \sqrt{x^2 + y^2}$ is the standard vector magnitude. Then the collection $\mathcal{A}$ is the set of all circular regions:

$$\mathcal{A} = \left\{C_{\mathbf{x}_0,r} \mid \mathbf{x}_0 \in \mathbb{R}^2 \text{ and } r \in \mathbb{R}_+\right\} .$$

Then let $\mathcal{C} = \left\{C_{(0,0),r} \mid r \in \mathbb{R}_+\right\}$ be the set of circles centered at the origin, which is a maximal simply ordered subset according to the example (and this is not difficult to show). Arbitrarily choose $X = C_{(0,0),1}$, that is the circular region of radius 1 centered at the origin, so that clearly $X \in \mathcal{C}$. Since the partial order in this is example is "is a proper subset of", define

$$\mathcal{B}' = \{Y \in \mathcal{A} \mid Y \subsetneq X \text{ or } X \subsetneq Y\}$$

and $\mathcal{B} = \mathcal{B}' \cup \{X\}$. The question is then whether $\mathcal{B} = \mathcal{C}$. We claim that, for this example, this is not the case.

Proof. Consider the set $C_{(1,0),2}$ and any $\mathbf{x} \in X = C_{(0,0),1}$ so that $|\mathbf{x} - (0,0)| < 1$. Then we have

$$|\mathbf{x} - (1,0)| \leq |\mathbf{x} - (0,0)| + |(0,0) - (1,0)| < 1 + |(-1,0)| = 1 + 1 = 2\,,$$

where we have utilized the ever-useful triangle inequality. Therefore $\mathbf{x} \in C_{(1,0),2}$ so that $X \subset C_{(1,0),2}$ since $\mathbf{x}$ was arbitrary. However, clearly the point $(1,0) \in C_{(1,0),2}$ but we have that $(1,0) \notin C_{(0,0),1} = X$ since $|(1,0) - (0,0)| = |(1,0)| = 1 \geq 1$. This shows that $X \subsetneq C_{(1,0),2}$ so that by definition $C_{(1,0),2} \in \mathcal{B}'$ and therefore $C_{(1,0),2} \in \mathcal{B} = \mathcal{B}' \cup \{X\}$. But clearly $C_{(1,0),2} \notin \mathcal{C}$ since it is not centered at the origin. This shows that $\mathcal{B} \neq \mathcal{C}$ as desired. $\square$

Hence it would seem that this method of attempting to define a maximal simply ordered subset containing X has failed in this example. It is easy to come up with an analogous counterexample that shows the same result of the other example of a maximal simply ordered subset of circles tangent to the y-axis at the origin.

Regarding Example 2, recall that the order $\prec$ is defined by

$$(x_0, y_0) \prec (x_1, y_1)$$

if $y_0 = y_1$ and $x_0 < x_1$ for (x_0, y_0) and (x_1, y_1) in $\mathbb{R}^2$. It is then claimed (which is again easy to show) that maximal simply ordered subsets are horizontal lines in the plane, that is sets

$$L_{y_0} = \{(x,y) \in \mathbb{R}^2 \mid y = y_0\}$$

for some $y_0 \in \mathbb{R}$. So consider any such $y_0 \in \mathbb{R}$ and let $\mathbf{x} = (0, y_0)$. Now define

$$B' = \{\mathbf{y} \in \mathbb{R}^2 \mid \mathbf{x} \prec \mathbf{y} \text{ or } \mathbf{y} \prec \mathbf{x}\}$$

and $B = B' \cup \{\mathbf{x}\}$. In contrast to Example 1, we here claim that $B = L_{y_0}$, which is to say that B *does* define the maximal simply ordered subset.

Proof. Consider any $(x,y) \in B = B' \cup \{\mathbf{x}\}$ so that either $(x,y) \in B'$ or $(x,y) = \mathbf{x}$. Clearly if $(x,y) = \mathbf{x} = (0, y_0)$ then $(x,y) \in L_{y_0}$ since $y = y_0$. On the other hand, if $(x,y) \in B'$ then $(x,y) \prec \mathbf{x}$ or $\mathbf{x} \prec (x,y)$. In the former case we have that $(x,y) \prec \mathbf{x} = (0, y_0)$ so that, by definition $y = y_0$ and $x < 0$. Clearly then $(x,y) = (x, y_0) \in L_{y_0}$ by definition. In the latter case we also have $y = y_0$ (though this time $0 < x$) so that again $(x,y) \in L_{y_0}$. Since (x,y) was arbitrary, this shows that $B \subset L_{y_0}$.

Now consider any $(x,y) \in L_{y_0}$ so that $y = y_0$. If $x = 0$ then $(x,y) = (0, y_0) = \mathbf{x}$ so that obviously $(x,y) \in \{\mathbf{x}\}$. If $0 < x$ then $(x,y) = (x, y_0) \prec (0, y_0) = \mathbf{x}$ so that $(x,y) \in B'$. Similarly, if $x < 0$, then $\mathbf{x} = (0, y_0) \prec (x, y_0) = (x,y)$ so that again $(x,y) \in B'$. Hence in all cases either $(x,y) \in B'$ or $(x,y) \in \{\mathbf{x}\}$ so that $(x,y) \in B' \cup \{\mathbf{x}\} = B$. This shows that $L_{y_0} \subset B$ since again (x,y) was arbitrary.

Thus we have shown that $B = L_{y_0}$ as desired. $\square$

So it would seem that, in this example, this naive technique does work!

Exercise 11.4

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:

First define

$$A = \left\{ (x, y) \in \mathbb{R}^2 \mid y = x^3 \right\} .$$

We show that it is a maximal simply ordered subset of $\mathbb{R}^2$.

Proof. First we show that A is simply ordered by $\prec$. Consider distinct (x_0, y_0) and (x_1, y_1) in A so that $y_0 = x_0^3$ and $y_1 = x_1^3$. Since they are distinct, it has to be that $x_0 \neq x_1$ or $y_0 \neq y_1$. The latter case actually implies the former since the function $f(x) = x^3$ is a well-defined function. Hence we can assume that $x_0 \neq x_1$, from which we can also assume without loss of generality that $x_0 < x_1$. Since $f(x) = x^3$ is also a monotonically increasing function (which is easy to show), it then follows that $y_0 = x_0^3 < x_1^3 = y_1$. Thus we have that $x_0 < x_1$ and $y_0 \leq y_1$ so that $(x_0, y_0) \prec (x_1, y_1)$ by definition. Since (x_0, y_0) and (x_1, y_1) were arbitrary, this shows that $\prec$ is a simple order on A.

To show that it is maximal suppose that B is any proper superset of A so that there is an $(x, y) \in B$ where $(x, y) \notin A$. Therefore clearly $y \neq x^3$ by definition. Now let $z = x^3$ so that $y \neq x^3 = z$ but $(x, z) \in A$. Clearly it is not true that $x < x$ so that it can neither be that $(x, y) \prec (x, z)$ nor $(x, z) \prec (z, y)$. Hence (x, y) and (x, z) are incomparable in $\prec$. This shows that B is not simply ordered and thus that A is maximal since B was an arbitrary superset. $\square$

Now redefine

$$A = \left\{ (x, y) \in \mathbb{R}^2 \mid y = 2 \right\} ,$$

which we also show is a maximal simply ordered subset of $\mathbb{R}^2$.

Proof. To show that A is simply ordered consider distinct (x_0, y_0) and (x_1, y_1) in A so that $y_0 = y_1 = 2$. Since these points are distinct and $y_0 = y_2$ is must be that $x_0 \neq x_1$, from which we can assume that $x_0 < x_1$ without loss of generality. But then clearly it is true that $x_0 < x_1$ and $y_0 \leq y_1$ so that $(x_0, y_0) \prec (x_1, y_1)$. Since these points were arbitrary this shows that A is simply ordered by $\prec$.

To show that it is maximal suppose that B is any proper superset of A so that there is an $(x, y) \in B$ where $(x, y) \notin A$. It then follows that $y \neq 2$ so that the point $(x, 2) \in A$ but $(x, 2) \neq (x, y)$. Clearly it can be that neither $(x, 2) \prec (x, y)$ nor $(x, y) \prec (x, 2)$ since it is not true that $x < x$. Hence (x, y) and $(x, 2)$ are incomparable in $\prec$. This shows that B is not simply ordered by $\prec$. Since B was an arbitrary superset this shows that A is maximal. $\square$

Now let

$$A = \left\{ (x, y) \in \mathbb{R}^2 \mid y = x^2 \right\} .$$

We claim that this subset is not simply ordered by $\prec$ and therefore cannot be a maximal simply ordered subset.

Proof. Consider the clearly distinct points $(-1, 1)$ and $(0, 0)$. Clearly since $0 = 0^2$ and $1 = (-1)^2$ these are both in A. However, since $1 > 0$ it is not true that $-1 < 0$ and $1 \leq 0$, and therefore it is not true that $(-1, 1) \prec (0, 0)$. Similarly since $0 \geq -1$ it is not true that $0 < -1$ and $0 \leq 1$, and therefore it is not true that $(0, 0) \prec (-1, 1)$. Hence the two distinct points are both in A but are not comparable. This suffices to show that A is not simply ordered by $\prec$. $\square$

We now claim that the maximal simply ordered subsets of $\mathbb{R}^2$ as ordered by $\prec$ are exactly the collection of sets of the form

$$A_f = \left\{(x,y) \in \mathbb{R}^2 \mid y = f(x)\right\}$$

for some function $f : (a,b) \to \mathbb{R}$, where (a,b) is an open interval of $\mathbb{R}$, noting that it could be that $a = -\infty$ and/or $b = \infty$. The function f must also satisfy the following properties:

(i) It is non-decreasing. Recall that this means that $x < y$ implies that $f(x) \leq f(y)$ for any $x, y \in (a,b)$.

(ii) If $b < \infty$ then its image is unbounded above.

(iii) If $a > -\infty$ then its image is unbounded below.

Now, let $\mathcal{A}$ be the collection of all these subsets and let $\mathcal{B}$ denote the set of all maximal simply ordered subsets. We show that $\mathcal{A} = \mathcal{B}$.

Proof. ($\subset$) First consider any $A_f \in \mathcal{A}$ so that $f : (a,b) \to \mathbb{R}$ with the properties above for some open interval (a,b). To show that A_f is simply ordered by $\prec$ consider any distinct (x,y) and (x',y') in A_f so that $y = f(x)$ and $y' = f(x')$. Since these are distinct it follows that $x \neq x'$ or $f(x) = y \neq y' = f(x')$. In the latter case it also follows that $x \neq x'$ as well for otherwise f would not be a function. Hence we can, without loss of generality, assume that $x < x'$. Since f is non-decreasing it follows that also $y = f(x) \leq f(x') = y'$, and therefore by definition $(x,y) \prec (x',y')$. Since these elements of A_f were arbitrary, it follows that A_f is simply ordered by $\prec$.

To show that A_f is maximal consider any proper superset A of A_f so that there is an $(x,y) \in A$ where $(x,y) \notin A_f$. There are a few possible ways in which (x,y) can fail to be an element of A_f.

Case: $x \in (a,b)$. Then it must be that $y \neq f(x)$ since $(x,y) \notin A_f$. Since it is not true that $x < x$, it has to be that neither $(x,y) \prec (x, f(x))$ nor $(x, f(x)), (x,y)$. Hence (x,y) and $(x, f(x))$ are incomparable elements of A (noting that clearly $(x, (f(x)) \in A_f \subset A)$ so that A is not simply ordered by $\prec$.

Case: $x \geq b$. Note that this is only possible if $b < \infty$ so that $b \in \mathbb{R}$. Thus in this case we have that the image of f is unbounded above by property (ii). Hence there is a $y_u \in reals$ where $y' > y$ and y' is in the image of f. Thus there is also an $x' \in (a,b)$ where $y = f(x')$ so that $(x',y') \in A_f \subset A$. Now, we have $x' < b \leq x$ but $y' > y$ so that it is not true that $y' \leq y$, and hence it cannot be that $(x',y') \prec (x,y)$. Similarly is it is clearly not true that $x < x'$ so that it cannot be that $(x,y) \prec (x',y')$ either. This shows that (x,y) and (x',y') are incomparable elements of A so that A is not simply ordered.

Case: $x \leq a$. An argument analogous to the previous case shows that $a > -\infty$ so that the image of f is unbounded below. From this it follows again that A is not simply ordered.

Thus in all cases A is not simply ordered so that A_f is a maximal simply ordered subset of $\mathbb{R}^2$ since A was an arbitrary proper superset. This shows that $A_f \in \mathcal{B}$ so that $\mathcal{A} \subset \mathcal{B}$ since A_f was arbitrary.

($\supset$) Now consider any $B \in \mathcal{B}$ so that B is a maximal simply ordered set by $\prec$. Define

$$X = \{x \in \mathbb{R} \mid (x,y) \in B \text{ for some } y \in \mathbb{R}\} .$$

We prove that B has the following properties:

(1) If (x_0, y_0) and (x_1, y_1) are in B and $x_0 < x_1$ then $y_0 \leq y_1$.

(2) For every $x \in X$ there is a unique $y \in \mathbb{R}$ where $(x,y) \in B$.

To show show (1) consider (x_0, y_0) and (x_1, y_1) in B and suppose that $x_0 < x_1$. Since B is simply ordered, it must be that either $(x_0, y_0) \prec (x_1, y_1)$ or $(x_1, y_1) \prec (x_0, y_0)$. Since $x_0 < x_1$ it clearly must be that $(x_0, y_0) \prec (x_1, y_1)$ and hence also $y_0 \le y_1$.

To show (2) consider any $x \in X$. Clearly there is a $y \in \mathbb{R}$ where $(x, y) \in B$ by the definition of X. To show that this y is unique, suppose that (x, y_0) and (x, y_1) are both in B but that $y_0 \neq y_1$ so that (x, y_0) and (x, y_1) are distinct. Since B is simply ordered they must be comparable in $\prec$ but they clearly cannot be since it is not true that $x < x$. As this is a contradiction, it must be that $y_0 = y_1$.

With that out of the way, let b be the least upper bound of X if it is bounded above and $b = \infty$ otherwise. Similarly let a be the greatest lower bound if X is bounded below and $a = -\infty$ otherwise. Now we claim that X is equal to the open interval (a, b).

So consider any $x \in X$ so that then clearly $a \le x \le b$ since a and b are lower and upper bounds of X, respectively. Clearly if $b = \infty$ then it cannot be that $x = b$ (since $x \in \mathbb{R}$) so assume that $b \in \mathbb{R}$ and $x = b$. Then $b = x \in X$ so that by property (2) there is a unique $y \in \mathbb{R}$ where $(b, y) \in B$. Clearly then $(b+1, y) \notin B$, since $b+1 \notin X$, so that the set $B' = B \cup \{(b+1, y)\}$ is a proper superset of B. Now consider any $(x', y') \in B$ so that clearly $x' \in X$ and hence $x' \le b < b+1$. By property (1) above it also follows that $y' \le y$, and so we have that $(x', y') \prec (b+1, y)$. Since (x', y') was arbitrary, this shows that $(b+1, y)$ is comparable to every element of B and hence B' is simply ordered by $\prec$. But this is not possible since B is maximal and B' is a proper superset. Hence it must be that $x \neq b$. An analogous argument shows that $x \neq a$ as well and hence $a < x < b$. Since x was arbitrary this shows that $X \subset (a, b)$.

Now consider any $x \in (a, b)$ so that $a < x < b$. Since b is the least upper bound of X, it has to be that x is not an upper of X so that there is an $x_g \in X$ where $x < x_g < b$ (clearly the existence of x_g also follows when $b = \infty$ since then X is unbounded above). Clearly then there is also a $y_g \in \mathbb{R}$ where $(x_g, y_g) \in B$. It then follows that the set $Y_g = \{y \in \mathbb{R} \mid (z, y) \in B \text{ for some } x < z < b\}$ is nonempty. By an analogous argument there is an $(x_l, y_l) \in B$ where $a < x_l < x$ so that the set $Y_l = \{y \in \mathbb{R} \mid (z, y) \in B \text{ for some } a < z < x\}$ is nonempty. Now, for any $y \in Y_g$, we have that $(z, y) \in B$ for some $x < z < b$. Therefore $x_l < x < z$ and by property (1) of B we have that $y_l \le y$. Since y was arbitrary this shows that y_l is a lower bound of Y_g and hence it has a greatest lower bound y_v. So suppose that $x \notin X$ so that there is not a $y \in \mathbb{R}$ where $(x, y) \in B$. Then we have that the set $B \cup \{(x, y_v)\}$ is a proper superset of B. However, consider any $(x', y') \in B$ so that $x' \in X$ but $x' \neq x$.

Case: $x' < x$. Then it has to be that $a < x' < x$ so that $y' \in Y_l$. Then, for any $y \in Y_g$ we again have that $(z, y) \in B$ for some $x < z < b$. Hence $x' < x < z$ so that $y' \le y$ by property (1) since $(x', y') \in B$ and $(z, y) \in B$. Since y was arbitrary, this shows that y' is a lower bound of Y_g. Since y_v is the greatest lower bound of Y_g, we have that $y' \le y_v$. Then clearly $(x', y') \prec (x, y_v)$ since also $x' < x$.

Case: $x' > x$. Then it has to be that $x < x' < b$ so that $y' \in Y_g$. It then follows that $y_v \le y'$ since y_v is the greatest lower bound of Y_g. Hence we have that $(x, y_v) \prec (x', y')$ since $x < x'$ as well.

Therefore in all cases we have that (x, y_v) and (x', y') are comparable in $\prec$. Since (x', y') was arbitrary, this clearly shows that $B \cup \{(x, y_v)\}$ is simply ordered. But this cannot be possible since it is a proper superset and B is maximal! So it has to be that in fact there *is* a $y \in \mathbb{R}$ where $(x, y) \in B$, and hence $x \in X$. Since $x \in (a, b)$ was arbitrary, this shows that $(a, b) \subset X$. This completes the rather long proof that $X = (a, b)$.

Now, by property (2) there is a unique $y \in \mathbb{R}$ for every $x \in X = (a, b)$ where $(x, y) \in B$. So we define a function $f : (a, b) \to \mathbb{R}$ by simply setting $f(x) = y$. Clearly based on the way this function is defined and the fact that $(a, b) = X$ we have that $B = A_f$. We must now show that f has the properties (i) through (iii) above.

Property (i) follows almost immediately from property (1) of B. To see this, consider any $x, y \in (a, b)$ where $x < y$. Then $(x, f(x))$ and $(y, f(y))$ are in B and hence $f(x) \leq f(y)$ by property (1). For property (ii) suppose that $b < \infty$ but that the image of f is bounded above. Hence it image has an upper bound, say $y_u \in \mathbb{R}$, so that clearly $B \cup \{(b+1, y_u)\}$ is a proper superset of B. So consider any $(x, y) \in B$ so that $y = f(x)$ for some $x \in (a, b)$. Then clearly $f(x)$ is in the image of f so that $y = f(x) \leq y_u$ since y_u is an upper bound of the image. Since also we must have $x < b < b+1$, it follows that $(x, y) \prec (b+1, y_u)$. Since $(x, y) \in B$ was arbitrary, this shows that $B \cup \{(b+1, y_u)\}$ is simply ordered, which cannot be possible since it is a proper superset and B is maximal. So it has to be that in fact the image of f is unbounded above when $b < \infty$, which shows property (ii). An analogous argument shows property (iii).

Since f has all of the required properties and $B = A_f$, this shows that $B \in \mathcal{A}$. Clearly then $\mathcal{B} \subset \mathcal{A}$ since B was arbitrary. This shows that $\mathcal{A} = \mathcal{B}$ as desired. □

Lastly, note that the example curves $y = x^3$ and $y = 2$ are clearly in $\mathcal{A} = \mathcal{B}$ since they are non-decreasing functions on $\mathbb{R}$, ($\mathbb{R}$ being the same as the open interval $(-\infty, \infty)$), while the curve $y = x^2$ is not since it is decreasing when $x < 0$.

Exercise 11.5

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 $\mathcal{A}$ has an element that is properly contained in no other element of $\mathcal{A}$.

Solution:

Proof. First, we know that $\subsetneq$ is a strict partial order on $\mathcal{A}$, which is trivial to show. So consider any simply ordered subset $\mathcal{B}$ of $\mathcal{A}$ and let $A = \bigcup \mathcal{B}$ so that we know that $A \in \mathcal{A}$. Clearly for any set $B \in \mathcal{B}$ we have that $B \subset \bigcup \mathcal{B} = A$ so that, since B was arbitrary, A is an upper bound of $\mathcal{B}$ in the strict partial order $\subsetneq$. Since $\mathcal{B}$ was arbitrary, this shows the hypothesis of Zorn's Lemma so that $\mathcal{A}$ has a maximal element A. Then clearly A is not properly contained in any other element of $\mathcal{A}$. □

Exercise 11.6

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 the 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 properly contained in no other element of $\mathcal{A}$.

Solution:

Proof. Suppose that $\mathcal{A}$ is a collection of sets of finite type. Let $\mathcal{B}$ be a subcollection of $\mathcal{A}$ that is simply ordered by $\subsetneq$. Consider next any finite subset B of $\bigcup \mathcal{B}$. Then, for every $b \in B$, $b \in \bigcup \mathcal{B}$ so that we can choose a set $B_b \in \mathcal{B}$ such that $b \in B_b$. Note that this does not require the choice axiom since we need to make only a finite number of choices. Then the set $\mathcal{B}' = \{B_b \mid b \in B\}$ is clearly a finite set of elements of $\mathcal{B}$. Since $\mathcal{B}$ is simply ordered by $\subsetneq$, it follows that $\mathcal{B}'$ is as well and so has a largest element C since it is finite.

Hence, for any $b \in B$, we have that $b \in B_b \subset C$ so that $b \in C$, and so B is a finite subset of C. Since $C \in \mathcal{B}$ and $\mathcal{B} \subset \mathcal{A}$, clearly $C \in \mathcal{A}$. Since $\mathcal{A}$ is of finite type and B is a finite subset of C, it follows that $B \in \mathcal{A}$ also. Since B was an arbitrary finite subset of $\bigcup \mathcal{B}$, it then follows that $\bigcup \mathcal{B}$ is also in $\mathcal{A}$ since it is of finite type. It then follows from the Kuratowski lemma (Exercise 11.5) that $\mathcal{A}$ has an element that is properly contained in no other element of $\mathcal{A}$ as desired. $\square$

Exercise 11.7

Show that the Tukey lemma implies the Hausdorff maximum principle. [Hint: If $\prec$ is a strict partial order on A, let $\mathcal{A}$ be the collection of all subsets of A that are simply ordered by $\prec$. Show that $\mathcal{A}$ is of finite type.]

Solution:

Proof. Following the hint, suppose that the set A has strict partial order $\prec$ and let $\mathcal{A}$ be the collection of all subsets of A that are simply ordered by $\prec$. We show that $\mathcal{A}$ has finite type, i.e. that a subset $B \subset A$ is in $\mathcal{A}$ if and only if every finite subset of B is.

($\Rightarrow$) Suppose that $B \subset A$ is in $\mathcal{A}$ so that it is simply ordered by $\prec$. Clearly any finite subset of B is also simply ordered by $\prec$ so that it is also in $\mathcal{A}$, which shows the result.

($\Leftarrow$) Now suppose that $B \subset A$ and that every finite subset of B is in $\mathcal{A}$. Now consider two distinct element x and y of B. Clearly then the set $\{x, y\}$ is a finite subset of B and hence is in $\mathcal{A}$. Then this means that $\{x, y\}$ is simply ordered by $\prec$ so that clearly x and y are comparable. Since x and y were arbitrary this shows that B is simply ordered by $\prec$ and hence $B \in \mathcal{A}$.

We have thus shown that $\mathcal{A}$ is of finite type so that it has a set C such that is properly contained in no other element of $\mathcal{A}$. Since $C \in \mathcal{A}$, it is simply ordered by $\prec$. It is also maximal since, if D is any proper superset of C then it cannot be that D is simply ordered for then we would have $D \in \mathcal{A}$ and $C \subsetneq D$, which would contradict the definition of C. Hence C is the maximal simply ordered subset of A that shows the maximum principle. $\square$

Exercise 11.8

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 that *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 the zero vector is the trivial one having all coefficients zero. If A is independent and if every vector in V belongs to the span of A, then A is a *basis* for 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 of all independent sets in V has a maximal element.
(c) Show that V has a basis.

Solution:
(a)

Proof. We show this by contradiction. Suppose that A is independent and $v \in V$ does not belong to the span of A. Also let $B = A \cup \{v\}$ and suppose that B is *not* independent. Then

$$\sum_{i=1}^{n} \beta_i b_i = 0$$

for some nonzero coefficients β_i, where each b_i is in B. Now, it must be that one of the b_i vectors is v and the rest in A since otherwise they would all be in A and then A would not be independent. Hence this can be expressed as

$$\sum_{i=1}^{n-1} \alpha_i a_i + \gamma v = 0$$

for nonzero coefficients α_i and γ and vectors $a_i \in A$. However clearly then we would have

$$v = -\frac{1}{\gamma} \sum_{i=1}^{n-1} \alpha_1 a_i = \sum_{i=1}^{n-1} \left(\frac{-\alpha_i}{\gamma} \right) a_i$$

so that v is a linear combination of vectors in A and hence is in the span of A. This is a contradiction so that it must be that in fact $B = A \cup \{v\}$ is independent as desired. $\square$

(b)

Proof. Let $\mathcal{A}$ be the collection of all independent sets in V. We know that $\subsetneq$ is a strict partial order on $\mathcal{A}$. Now let $\mathcal{B}$ be any subset of $\mathcal{A}$ that is simply ordered by $\subsetneq$. We claim that $\bigcup \mathcal{B}$ is an upper bound of $\mathcal{B}$ that is in $\mathcal{A}$. So first consider any $B \in \mathcal{B}$ and any $b \in B$ so that clearly then $b \in \bigcup \mathcal{B}$. Hence $B \subset \bigcup \mathcal{B}$ since b was arbitrary. Since $B \in \mathcal{B}$ was arbitrary, this shows that $\bigcup \mathcal{B}$ is an upper bound of $\mathcal{B}$ by $\subsetneq$.

Next we show that $\bigcup \mathcal{B}$ is also in $\mathcal{A}$. To this end consider any finite set B of elements of $\bigcup \mathcal{B}$ so that B is a set of vectors in V. Now, for each $b \in B$ we have that $b \in \bigcup \mathcal{B}$ so that we can choose any set $B_b \in \mathcal{B}$ where $b \in B_b$. Note that this does not require the axiom of choice since B is finite. Then, since each B_b is in $\mathcal{B}$, which is simply ordered by $\subsetneq$ and $\{B_b \mid b \in B\}$ is finite, it follows that it has a largest element C so that $B_b \subset C$ for any $b \in B$. Hence $B \subset C$ since each $b \in B_b$ and $B_b \subset C$. Also $C \in \mathcal{A}$ since $C \in \mathcal{B}$ and $\mathcal{B} \subset \mathcal{A}$ so that C is independent. Hence the only linear combination of the vectors in B that is the zero vector must have all zero coefficients since they are all in the independent set C. Since B was an arbitrary set of vectors in $\bigcup \mathcal{B}$, this shows that $\bigcup \mathcal{B}$ is independent and therefore in $\mathcal{A}$.

Since $\mathcal{B}$ was an arbitrary simply ordered subset of $\mathcal{A}$, it follows that every such subset has an upper bound in $\mathcal{A}$. Thus by Zorn's Lemma $\mathcal{A}$ has a maximal element as desired. $\square$

(c)

Proof. Again let $\mathcal{A}$ be the collection of all independent sets in V, which we know has a maximal element A from part (b). We claim that A is a basis for V. Suppose to the contrary that it is not so that, since we know that A is independent (since it is in $\mathcal{A}$), there must be a vector $v \in V$ that is not in the span of A. Then by part (a) we have that $A \cup \{v\}$ is also independent and so in $\mathcal{A}$. We also have that $v \notin A$ since otherwise it would clearly be in the span of A. Hence $A \subsetneq A \cup \{v\}$. However, this contradicts the fact that A is a maximal element of $\mathcal{A}$, so that it must be that in fact A is a basis for V as desired. $\square$

Supplementary Exercises: Well-Ordering

Exercise WO.1

Theorem (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 is a unique

$h : J \to C$ such that $h(\alpha) = \rho(h \upharpoonright S_\alpha)$ for each $\alpha \in J$. [Hint: Follow the pattern outlined in Exercise 10 of §10.]

Solution:

Following the hint, we follow the pattern of Exercise 10.10. In what follows denote by $(*)$ the property

$$h(\alpha) = \rho(h \upharpoonright S_\alpha)$$

for a function h from J or a section of J to C.

Lemma WO.1.1. *If h and k map sections of J, or all of J, into C and satisfy $(*)$ for all x in their respective domains, then $h(x) = k(x)$ for all x in both domains.*

Proof. First suppose that the domains of h and k are sets H and K where each is either a section of J or J itself. Since this is the case, we can assume without loss of generality that $H \subset K$ and so H is exactly the domain common to both h and k. Now suppose that the hypothesis we are trying to prove is *not* true so that there is an x in both domains (i.e. $x \in H$) where $h(x) \neq k(x)$. We can also assume that x is the smallest such element since $H \subset J$ and J is well-ordered. It then clearly follows that $S_x \subset H$ is a section of J and that $h(y) = k(y)$ for all $y \in S_x$. From this we clearly have that $h \upharpoonright S_x = k \upharpoonright S_x$ so that

$$h(x) = \rho(h \upharpoonright S_x) = \rho(k \upharpoonright S_x) = k(x)$$

since both h and k satisfy $(*)$ and x is in the domain of both. This contradicts the supposition that $h(x) \neq k(x)$ so that it must be that no such x exists and hence h and k are the same in their common domain as desired. $\square$

Lemma WO.1.2. *If there exists a function $h : S_\alpha \to C$ satisfying $(*)$, then there exists a function $k : S_\alpha \cup \{\alpha\} \to C$ satisfying $(*)$.*

Proof. Suppose that $h : S_\alpha \to C$ is such a function satisfying $(*)$. Now let $\bar{S}_\alpha = S_\alpha \cup \{\alpha\}$ and we define $k : \bar{S}_\alpha \to C$ as follows. For any $x \in \bar{S}_\alpha$ set

$$k(x) = \begin{cases} h(x) & x \in S_\alpha \\ \rho(h) & x = \alpha \,. \end{cases}$$

We note that clearly S_α and $\{\alpha\}$ are disjoint so that this is unambiguous. We also note that h is a function from a section of J to C so that $h \in \mathcal{F}$ and $\rho(h) \in C$ is therefore defined.

Now we show that k satisfies $(*)$. First, clearly $h \upharpoonright S_x = k \upharpoonright S_x$ for any $x \leq \alpha$ since $k(y) = h(y)$ by definition for any $y \in S_x \subset S_\alpha$. Now consider any $x \in \bar{S}_\alpha$. If $x = \alpha$ then by definition we have

$$k(x) = \rho(h) = \rho(h \upharpoonright S_\alpha) = \rho(k \upharpoonright S_\alpha) = \rho(k \upharpoonright S_x)$$

since clearly $h = h \upharpoonright S_\alpha$ since S_α is the domain of h. On the other hand, if $x \in S_\alpha$ then $x < \alpha$ so that

$$k(x) = h(x) = \rho(h \upharpoonright S_x) = \rho(k \upharpoonright S_x)$$

since h satisfies $(*)$. Therefore, since x was arbitrary, this shows that k also satisfies $(*)$. $\square$

Lemma WO.1.3. *If $K \subset J$ and for all $\alpha \in K$ there exists a function $h_\alpha : S_\alpha \to C$ satisfying $(*)$, then there exists a function*

$$k : \bigcup_{\alpha \in K} S_\alpha \to C$$

satisfying $(*)$.

Proof. Let

$$k = \bigcup_{\alpha \in K} h_\alpha \,,$$

which we claim is the function we seek.

First we show that k is actually a function from $\bigcup_{\alpha \in K} S_\alpha$ to C. So consider any x in the domain of k. Suppose that (x,a) and (x,b) are both in k so that there are α and β in K where $(x,a) \in h_\alpha$ and $(x,b) \in h_\beta$. Since h_α and h_β both satisfy $(*)$, it follows from Lemma WO.1.1 that $a = h_\alpha(x) = h_\beta(x) = b$ since clearly x is in the domain of both. This shows that k is indeed a function since (x,a) and (x,b) were arbitrary. Also clearly the domain of k is $\bigcup_{\alpha \in K} S_\alpha$ since, for any $x \in \bigcup_{\alpha \in K} S_\alpha$, we have that there is an $\alpha \in K$ where $x \in S_\alpha$. Hence x is in the domain of h_α and so in the domain of k. In the other direction, clearly if x is in the domain of k then it is in the domain of h_α for some $\alpha \in K$. Since this domain is S_α, clearly $x \in \bigcup_{\alpha \in K} S_\alpha$. Lastly, obviously the range of k can be C since this is the range of every h_α.

Now we show that k satisfies $(*)$. So consider any $x \in \bigcup_{\alpha \in K} S_\alpha$ so that $x \in S_\alpha$ for some $\alpha \in K$. Clearly we have that $k(y) = h_\alpha(y)$ for every $y \in S_\alpha$ since $h_\alpha \subset k$. It then immediately follows that $k(x) = h(x)$ and $k \restriction S_x = h_\alpha \restriction S_x$ since $S_x \subset S_\alpha$. Then, since h_α satisfies $(*)$, we have

$$k(x) = h_\alpha(x) = \rho(h_\alpha \restriction S_x) = \rho(k \restriction S_x)\,.$$

Since x was arbitrary, this shows that k satisfies $(*)$ as desired. □

Lemma WO.1.4. *For every $\beta \in J$, there exists a function $h_\beta : S_\beta \to C$ satisfying $(*)$.*

Proof. We show this by transfinite induction. So consider any $\beta \in J$ and suppose that, for every $x \in S_\beta$, there is a function $h_x : S_x \to C$ satisfying $(*)$. Now, if β has an immediate predecessor α then we claim that $S_\beta = S_\alpha \cup \{\alpha\}$. First if $x \in S_\beta$ then $x < \beta$ so that $x \leq \alpha$ since α is the immediate predecessor of β. If $x < \alpha$ then $x \in S_\alpha$ and if $x = \alpha$ then $x \in \{\alpha\}$. Hence in either case we have that $x \in S_\alpha \cup \{\alpha\}$. Now suppose that $x \in S_\alpha \cup \{\alpha\}$. If $x \in S_\alpha$ then $x < \alpha < \beta$ so that $s \in S_\beta$. On the other hand if $x \in \{\alpha\}$ then $x = \alpha < \beta$ so that again $x \in S_\beta$. Thus we have shown that $S_\beta \subset S_\alpha \cup \{\alpha\}$ and $S_\alpha \cup \{\alpha\} \subset S_\beta$ so that $S_\beta = S_\alpha \cup \{\alpha\}$. Since $\alpha \in S_\beta$ it follows that there is an $h_\alpha : S_\alpha \to C$ that satisfies $(*)$. Then, by Lemma WO.1.2, we have that there is an $h_\beta : S_\beta = S_\alpha \cup \{\alpha\} \to C$ that also satisfies $(*)$.

If β does not have an immediate predecessor then we claim that $S_\beta = \bigcup_{\gamma<\beta} S_\gamma$. So consider any $x \in S_\beta$ so that $x < \beta$. Since x cannot be the immediate predecessor of β, there must be an α where $x < \alpha < \beta$. Then $x \in S_\alpha$ so that, since $\alpha < \beta$, clearly $x \in \bigcup_{\gamma<\beta} S_\gamma$. Now suppose that $x \in \bigcup_{\gamma<\beta} S_\gamma$ so that there is an $\alpha < \beta$ where $x \in S_\alpha$. Then clearly $x < \alpha < \beta$ so that also $x \in S_\beta$. Thus we have shown that $S_\beta \subset \bigcup_{\gamma<\beta} S_\gamma$ and $\bigcup_{\gamma<\beta} S_\gamma \subset S_\beta$ so that $S_\beta = \bigcup_{\gamma<\beta} S_\gamma$. Now, clearly S_β is a subset of J where there is an $h_x : S_x \to C$ satisfying $(*)$ for every $x \in S_\beta$. Then it follows from Lemma WO.1.3 that there is a function h_β from $\bigcup_{\gamma<S_\beta} S_\gamma = \bigcup_{\gamma<\beta} S_\gamma = S_\beta$ to C that satisfies $(*)$.

Therefore, in either case, we have shown that there is an $h_\beta : S_\beta \to C$ that satisfies $(*)$. The desired result then follows by transfinite induction. □

Main Problem.

Proof. First suppose that J has no largest element. Then we claim that $J = \bigcup_{\alpha \in J} S_\alpha$. For any $x \in J$ there must be a $y \in J$ where $x < y$ since x cannot be the largest element of J. Hence $x \in S_y$ so that also clearly $\bigcup_{\alpha \in J} S_\alpha$. Then, for any $x \in \bigcup_{\alpha \in J} S_\alpha$, there is an $\alpha \in J$ where $x \in S_\alpha$. Clearly $S_\alpha \subset J$ so that $x \in J$ also. Hence $J \subset \bigcup_{\alpha \in J} S_\alpha$ and $\bigcup_{\alpha \in J} S_\alpha \subset J$ so that $J = \bigcup_{\alpha \in J} S_\alpha$. Since we know from Lemma WO.1.4 that there is an $h_\alpha : S_\alpha \to C$ that satisfies $(*)$ for every $\alpha \in J$, it follows from Lemma WO.1.3 that there is a function h from $\bigcup_{\alpha \in J} S_\alpha = J$ to C that satisfies $(*)$.

If J does have a largest element β then clearly $J = S_\beta \cup \{\beta\}$. Since we know that there is an $h_\beta : S_\beta \to C$ that satisfies $(*)$ by Lemma WO.1.4, it follows from Lemma WO.1.2 that there is a function h from $S_\beta \cup \{\beta\} = J$ to C that satisfies $(*)$. Hence the desired function h exists in both cases. Lemma WO.1.1 also clearly shows that this function is unique. □

Exercise WO.2

(a) Let J and E be well-ordered sets; let $h : J \to E$. Show that the following statements are equivalent:

  (i) h is order preserving and its image is E or a section of E.

  (ii) $h(\alpha) = \text{smallest } [E - h(S_\alpha)]$ for all α.

  [Hint: Show that each of these conditions implies that $h(S_\alpha)$ is a section of E; conclude that it must be the section by $h(\alpha)$.]

(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. [Hint: Given J, there is a most one order preserving map of J into E whose image is E or a section of E.]

Solution:

(a)

Proof. First, for any $\alpha \in J$ and $\beta \in E$, let S_α denote the section of J by α, and T_β denote the section of E by β. To avoid ambiguity, also suppose that $<$ is the well-order on J and $\prec$ is the well-order on E. We show that each of these conditions are equivalent to the condition that $h(S_\alpha) = T_{h(\alpha)}$ for every $\alpha \in J$. Call this condition (iii). This of course also shows that the conditions are equivalent to each other.

First we show that (i) implies (iii). So suppose that h is order preserving and its image is E or a section of E. Consider any $\alpha \in J$ and any $y \in h(S_\alpha)$ so that there is an $x \in S_\alpha$ where $y = h(x)$. Then $x < \alpha$ and $y = h(x) \prec h(\alpha)$ since h preserves order. Therefore $y \in T_{h(\alpha)}$ so that $h(S_\alpha) \subset T_{h(\alpha)}$ since y was arbitrary. Now consider $y \in T_{h(\alpha)}$ so that $y \prec h(\alpha)$. Since also clearly $y \in E$ (since $T_{h(\alpha)} \subset E$), y is in the image of h if its image is all of E. If the image of h is some section of E, say T_β, then clearly $h(\alpha) \in T_\beta$ since $h(\alpha)$ is obviously in the image of h. Hence we have $y \prec h(\alpha) \prec \beta$ so that $y \in T_\beta$ and hence in the image of h. Since y is in the image of h in either case, there is an $x \in J$ such that $y = h(x)$. Then $h(x) = y \prec h(\alpha)$ so that $x < \alpha$ since h preserves order. Hence $x \in S_\alpha$ so that $y \in h(S_\alpha)$ since $y = h(x)$. This shows that $T_{h(\alpha)} \subset h(S_\alpha)$ since y was arbitrary. Therefore $h(S_\alpha) = T_{h(\alpha)}$ so that condition (iii) is true since α was arbitrary.

Next we show that (iii) implies (i). So suppose that $h(S_\alpha) = T_{h(\alpha)}$ for all $\alpha \in J$. First, it is easy to see that h preserves order since, if $x, y \in J$ where $x < y$, then we have that $x \in S_y$ so that clearly $h(x) \in h(S_y) = T_{h(y)}$, and hence $h(x) < h(y)$. To show that the image of h, i.e. $h(J)$, is either E or a section of E, consider the set $E - h(J)$.

Case: $E - h(J) = \varnothing$. Then clearly for any $y \in E$ we must have that $y \in h(J)$ since otherwise it would be that $y \in E - h(J)$. Thus $E \subset h(J)$ since y was arbitrary. Also clearly $h(J) \subset E$ since E is the range of h. This shows that $h(J) = E$.

Case: $E - h(J) \neq \varnothing$. Then clearly $E - h(J)$ is a nonempty subset of E so that it has a smallest element β since E is well-ordered, noting that clearly $\beta \notin h(J)$. We claim that $h(J) = T_\beta$. So consider any $y \in h(J)$ so that there is an $x \in J$ where $y = h(x)$. Suppose for a moment that $\beta \preceq y$. Now it cannot be that $\beta = y$ since $y \in h(J)$ but $\beta \notin h(J)$, and so $\beta \prec y$. But then $\beta \in T_y = T_{h(x)} = h(S_x)$ since $x \in J$. Then β is in the image of h since clearly $h(S_x) \subset h(J)$. As this contradicts the fact that $\beta \notin h(J)$, it must be that $y \prec \beta$ so that $y \in T_\beta$. This shows that $h(J) \subset T_\beta$ since y was arbitrary. Suppose now that $y \in T_\beta$ so that $y \prec \beta$. Since β is the smallest element of $E - h(J)$, it follows that $y \notin E - h(J)$. Since clearly $y \in E$ (since $T_\beta \subset E$), it must be that $y \in h(J)$. This shows that $T_\beta \subset h(J)$ since y was arbitrary. Hence we have shown that $h(J) = T_\beta$.

Therefore in every case either the image of h is E or a section of E as desired. This completes the proof of (i).

Now we show that (ii) implies (iii). So suppose that $h(\alpha)$ is the smallest element of $E - h(S_\alpha)$ for every $\alpha \in J$. First we show that h is injective. So consider any $x, y \in J$ where $x \neq y$. We can assume without loss of generality that $x < y$ so that $x \in S_y$ and hence $h(x) \in h(S_y)$. However since we have that $h(y)$ is the smallest element of $E - h(S_y)$, clearly $h(y) \notin h(S_y)$. Therefore we have that $h(x) \neq h(y)$ so that h is injective.

Now consider any $\alpha \in J$ so that clearly $h(\alpha)$ is the smallest element of $E - h(S_\alpha)$. Suppose that $y \in h(S_\alpha)$ so that there is an $x \in S_\alpha$ where $y = h(x)$, and therefore $x < \alpha$. Consider the possibility that $h(\alpha) \preccurlyeq h(x) = y$. It cannot be that $h(\alpha) = h(x) = y$ since $x \neq \alpha$ and h is injective, so it must be that $h(\alpha) \prec h(x)$. It then follows that $h(\alpha) \notin E - h(S_x)$ since $h(x)$ is the smallest element of $E - h(S_x)$. Thus $h(\alpha) \in h(S_x)$ since clearly $h(\alpha) \in E$. It then follows from the fact that h is injective that $\alpha \in S_x$ so that we have $\alpha < x < \alpha$, which is clearly a contradiction. So it must be that $y = h(x) \prec h(\alpha)$ so that $y \in T_{h(\alpha)}$. This shows that $h(S_\alpha) \subset T_{h(\alpha)}$ since y was arbitrary.

Now suppose that $y \in T_{h(\alpha)}$ so that $y \prec h(\alpha)$. Since $h(\alpha)$ is the smallest element of $E - h(S_\alpha)$, it follows that $y \notin E - h(S_\alpha)$. Since clearly $y \in E$, it must be that $y \in h(S_\alpha)$. This shows that $T_{h(\alpha)} \subset h(S_\alpha)$ since y was arbitrary, and hence $h(S_\alpha) = T_{h(\alpha)}$, which shows (iii) since α was arbitrary.

Lastly, we show that (iii) implies (ii). So suppose that $h(S_\alpha) = T_{h(\alpha)}$ for every $\alpha \in J$ and consider any such α. Clearly we have that $h(\alpha) \in E$ but $h(\alpha) \notin T_{h(\alpha)} = h(S_\alpha)$ so that $h(\alpha) \in E - h(S_\alpha)$. Suppose for the moment that $h(\alpha)$ is not the smallest element of $E - h(S_\alpha)$ so that there is a $\beta \in E - h(S_\alpha)$ where $\beta \prec h(\alpha)$. Then $\beta \in T_{h(\alpha)}$ so that it must be that $\beta \notin E - T_{h(\alpha)} = E - h(S_\alpha)$ since $h(S_\alpha) = T_{h(\alpha)}$. Clearly this is a contradiction so that it must be that $h(\alpha)$ really is the smallest element of $E - h(S_\alpha)$, which shows (ii) since α was arbitrary. □

Exercise WO.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. [Hint: Choose $e_0 \in E$. Define $h : J \to E$ by the recursion formula

$$h(\alpha) = \text{smallest } [E - h(S_\alpha)] \qquad \text{if} \qquad h(S_\alpha) \neq E\,,$$

and $h(\alpha) = e_0$ otherwise. Show that $h(\alpha) \leq k(\alpha)$ for all α; conclude that $h(S_\alpha) \neq E$ for all α.]

Solution:

Proof. First, if $E = \varnothing$ then it must be that $J = \varnothing$ as well so that they vacuously have the same order type. Otherwise, following the hint, choose $e_0 \in E$ and define $h : J \to E$ by

$$h(\alpha) = \text{smallest } [E - h(S_\alpha)] \qquad \text{if} \qquad h(S_\alpha) \neq E\,,$$

and $h(\alpha) = e_0$ otherwise, noting that this function is uniquely defined by the general principle of recursive definition (Exercise WO.1). We show that $h(\alpha) \leq k(\alpha)$ for all $\alpha \in J$ using transfinite induction (see Lemma 10.10.1). So consider $\alpha \in J$ and assume that $h(x) \leq k(x)$ for all $x \in S_\alpha$.

Since k preserves order we have that $h(x) \leq k(x) < k(\alpha)$ when $x < \alpha$. In particular, this means that $h(x) \neq k(\alpha)$ for all $x \in S_\alpha$ so that $k(\alpha) \in E - h(S_\alpha)$. Hence $E - h(S_\alpha)$ is not empty so that $h(S_\alpha) \neq E$. Thus $h(\alpha)$ is the smallest element of $E - h(S_\alpha)$ and so $h(\alpha) \leq k(\alpha)$ since $k(\alpha) \in E - h(S_\alpha)$. This completes the induction.

Therefore, for any $\alpha \in J$ and any $x < \alpha$ we have $h(x) \leq k(x) < k(\alpha)$ since k preserves order so that $h(x) \neq k(\alpha)$. As in the induction step above, it follows that $h(S_\alpha) \neq E$. Hence, since α was arbitrary,

$$h(\alpha) = \text{smallest } [E - h(S_\alpha)]$$

for all $\alpha \in J$. It then follows from Exercise WO.2 part (a) that h is order preserving and maps J onto E or a section of E. This clearly shows that J has the order type of E or a section of E as desired. □

Exercise WO.4

Use Exercises 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. [Hint: Form a well-ordered set containing both A and B, as in Exercise 8 of §10; then apply the preceding exercise.]

(b) Suppose that A and B are well-ordered sets that are uncountable, such that every section of A and B is countable. Show that A and B have the same order type.

Solution:

(a)

Proof. First, we can assume that A and B are disjoint since, if not, we can form $A' = \{(x, 1) \mid x \in A\}$ and $B' = \{(x, 2) \mid x \in B\}$, which clearly *are* disjoint and have the same order types as A and B if ordered in the same way. So let $\prec$ be the order on $A \cup B$ as in Exercise 10.8 with all the elements of A before the elements of B. From the exercise, we know that $A \cup B$ is well-ordered by $\prec$. Now, clearly the identity function i_B with $A \cup B$ as the range is an order-preserving function from B to $A \cup B$ so that B is the same order type as $A \cup B$ or a section of $A \cup B$ by Exercise WO.3.

If B has the same order type as $A \cup B$, then there is a an order preserving bijection $g : A \cup B \to B$. Let b be the smallest element of B so that $y = g(b) \in B$. Since b is the smallest element of B, clearly the section $S_b = \{x \in A \cup B \mid x \prec b\} = A$. Also clearly $g(A) = g(S_b) = S_y = \{x \in B \mid x < y\}$ so that A has the same order type as a section of B since g preserves order.

If B has the same order type as a section of $A \cup B$ then there is an order preserving bijection $f : B \to S_\alpha$ for some $\alpha \in A \cup B$. If $\alpha \in A$ then clearly S_α lies entirely in A and is a section of A so that B has the same order type as a section of A. So now suppose that $\alpha \in B$. If α is the smallest element of B then again it has to be that S_α lies in A and is in fact the entirety of A so that B and A have the same order type. If α is not the smallest element of B then S_α contains elements of both A and B. So let b be the smallest element of B so that $b \in S_\alpha$, and let $y \in B$ be such that $f(y) = b$, which exists since f is surjective. We also have that $S_b = A$ since b is the smallest element of B. It then follows that $f(S_y) = S_b = A$ since $f(y) = b$ so that A has the same order type as the section S_y of B since f preserves order.

Hence in all cases one of the desired results always follows. To show that exactly one of these is the case, note that if A and B have the same order type then clearly it cannot be that A has the same order type as a section of B since then B would also have the same order type is its own section, which would violate Exercise WO.2b. Similarly B cannot have the same order type as a section of A since then A would have the same order type as its own section. Now suppose that A has the same order type as a section S_b of B. Then A and B cannot have the same order type since then B would have the same order type as its section S_b. Also B cannot have the same order type as a section S_a of A since then the section S_b, and therefore A, would have the same order type as a smaller section of A. An analogous argument shows the result when B has the same order type as a section of A. □

(b)

Proof. Suppose that A has the same order type as a section of B. Then there would be a bijection from A, an uncountable set, to a section of B, which is countable. A similar contradiction arises if B were to have the same order type as a section of A. By part (a), the only remaining possibility is that A and B have the same order type as desired. □

Exercise WO.5

Let X be a set; let $\mathcal{A}$ be the 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:

(a)

Proof. For any $(A, <) \in \mathcal{A}$ we have that it is not equal to a section of itself since then it would then clearly have the same order type as its own section, which would violate Exercise WO.2b. Hence it is not true that $(A, <) \prec (A, <)$ by definition, which shows that $\prec$ is nonreflexive.

Now consider $(A, <)$, $(A', <')$, and $(A'', <'')$ in $\mathcal{A}$ where $(A, <) \prec (A', <')$ and $(A', <') \prec (A'', <'')$. Then $(A, <)$ is a section of $(A', <')$. Also $(A', <')$ is a section of $(A'', <'')$ so that clearly any section of $(A', <')$ is also a section of $(A'', <'')$. Since $(A, <)$ is such a section we have that $(A, <)$ is a section of $(A'', <'')$ so that $(A, <) \prec (A'', <'')$. This shows that $\prec$ is transitive.

This completes the proof that $\prec$ is a strict partial order. □

(b)

Proof. First we must show that B' is simply ordered by $<'$.

First consider any $(x, y) \in <'$ so that there is a $(B, <) \in \mathcal{B}$ where $(x, y) \in <$ and $x, y \in B$. Clearly then x and y are in the union B' so that $(x, y) \in B' \times B'$. This shows that $<' \subset B' \times B'$ so that $<'$ is a relation on B'.

Next consider any x and y in B' where $x \neq y$. Then there are well-ordered sets $(B_1, <_1)$ and $(B_2, <_2)$ in $\mathcal{B}$ where $x \in B_1$ and $y \in B_2$. Since $\mathcal{B}$ is simply ordered by $\prec$ we have that $(B_1, <_1) \prec (B_2, <_2)$

or $(B_2, <_2) \prec (B_1, <_1)$. Without loss of generality we can assume the former case (since otherwise we can just swap the roles of x and y). Then $(B_1 <_1)$ is a section of $(B_2, <_2)$ and is thus also a subset so that $x, y \in B_2$, It then follows that x and y are comparable by $<_2$ since $x \neq y$ and $<_2$ is a well-order and therefore a simple order. Thus (x, y) or (y, x) are in $<_2$. Since $<'$ is the union of all relations $<$ where $(B, <) \in \mathcal{B}$, clearly we have that (x, y) or (y, x) are in $<'$ since $<_2$ is such a relation. Thus shows that $<'$ has the comparability property.

Now consider any $x \in B'$ so that there is a $(B, <) \in \mathcal{B}$ where $x \in B$. Consider also any $(B'', <'') \in \mathcal{B}$. Then, since $\mathcal{B}$ is simply ordered, it follows that $(B, <)$ and $(B'', <'')$ are comparable in $\prec$. If $(B, <) \prec (B'', <'')$ then $(B, <)$ is a section of $(B'', <'')$ so that $x \in B''$ as well. Then it cannot be that $x <'' x$ since $<''$ is a simple order. If $(B'', <'') \prec (B, <)$ then $(B'', <'')$ is a section of $(B, <)$. If $x \in B''$ then again it cannot be that $x <'' x$ since $<''$ is a simple order. If $x \notin B''$ then $(x, x) \notin <''$ since it is a relation on B''. Thus in all cases and sub-cases it is not true that $x <'' x$ so that $x <' x$ does not hold since $<''$ was arbitrary and $<'$ is their union. This shows that $<'$ is nonreflexive.

Lastly, suppose that $x <' y$ and $y <' z$. Then it has to be that there is a $(B_1, <_1)$ and $(B_2, <_2)$ in $\mathcal{B}$ where $z <_1 y$ and $y <_2 z$. Then $(B_1, <_1)$ and $(B_2, <_2)$ are comparable in $\prec$ since $\mathcal{B}$ is simply ordered. Hence one is a section of the other so that, in either case, it follows that $x < y$ and $y < z$ where either $< = <_1$ or $< = <_2$. Then clearly $x < z$ since both $<_1$ and $<_2$ are transitive since they are simple orders. Thus $x <' z$ since $<'$ is the union of all the orders in $\mathcal{B}$ and $<$ is such an order. This shows that $<'$ is transitive.

This completes the proof that $<'$ is a simple order on B'. To show that it is a well-order, consider any nonempty subset $A \subset B'$. Then there is an $x \in A$ so that $x \in B'$ as well. It then follows that there is a $(B, <) \in \mathcal{B}$ where $x \in B$. Then clearly $B \cap A$ is a nonempty subset of B since $x \in B$ and $x \in A$. Let b be the $<$-smallest element in $B \cap A$, and we claim that this is the smallest element of A by $<'$. First, obviously $b \in A$ since $b \in B \cap A$. Next consider any $y \in A$ so that $y \in B'$ as well. Then there is a $(B'', <'') \in \mathcal{B}$ where $y \in B''$. Since $\mathcal{B}$ is simply ordered by $\prec$ we have that $(B, <)$ and $(B'', <'')$ are comparable. Hence $(B, <)$ is a section of $(B'', <'')$ or vice-versa.

In the first case we have that both b and y are in B''. If $y \in B$ then also $y \in B \cap A$ so that $b \leq y$ since it is the smallest element of $B \cap A$ by $<$. If $y \notin B$ then $b <'' y$ since B is a section of B'', and therefore $b \leq'' y$ is true. In the second case in which $(B'', <'')$ is a section of $(B, <)$ we have that both b and y are in B and hence in $B \cap A$. Then, again $b \leq y$ since b is the smallest element of $B \cap A$ by $<$. Hence in all cases either $b \leq y$ or $b \leq'' y$. Either way it follows that $b \leq' y$ as well since $<'$ is the union. This shows that b is the smallest element of A by $<'$ as desired. Since A was an arbitrary nonempty subset, this shows that B' is well-ordered by $<'$. □

Exercise WO.6

Use Exercises 1 and 5 to prove the following:

Theorem. The maximum principle is equivalent to the well-ordering theorem.

Solution:

Proof. First suppose that the maximum principle is true and let X be any set. Then let $\mathcal{A}$ be the collection of all pairs $(A, <)$, where $A \subset X$ and $<$ is a well-ordering of A as in Exercise WO.5. Define the relation $\prec$ on $\mathcal{A}$ also as in Exercise WO.5, i.e $(A, <) \prec (A', <')$ if $(A, <)$ is a section of $(A', <')$. It was then shown in that exercise that $\prec$ is a strict partial order on $\mathcal{A}$ so that, by the maximum principle, there is a maximal simply ordered subset $\mathcal{B} \subset \mathcal{A}$. Now let $(B', <')$ be the unions of the corresponding elements of $\mathcal{B}$ so that we know that $<'$ well-orders B' by part (b) of Exercise WO.5.

We claim that $B' = X$. Suppose that this is not the case so that there is a $x \in X$ where $x \notin B'$ (since we know that $B' \subset X$). Then define $B'' = B' \cup \{x\}$ and the relation $<'' = <' \cup \{(b', x) \mid b' \in B'\}$. It is then easy to see (and trivial but tedious to show) that B'' is well-ordered by $<''$. Also, clearly B' is the section of B'' by x so that, for any $B \in \mathcal{B}$, we have $B \preceq B' \prec B''$. Since B was arbitrary, this shows that the set $\mathcal{B} \cup \{B''\}$ is simply ordered by $\prec$ and is a subset of $\mathcal{A}$. Since $x \notin B'$ we have that $x \notin B$ for any $B \in \mathcal{B}$ (since B' is their union) so that $B'' \neq B$ since $x \in B$. It follows that $\mathcal{B} \subsetneq \mathcal{B} \cup \{B''\}$, but this contradicts the maximality of $\mathcal{B}$! So it has to be that in fact $B' = X$ itself so that $<'$ is a well-ordering of X. Since X was an arbitrary set, this shows the well-ordering theorem.

Now suppose the well-ordering theorem and that A is a set with strict partial ordering $\prec$. Then we know that A has a well-ordering, say $<$. Now, for any function f from a section S_x (by $<$) to $\mathcal{P}(A)$, define

$$\rho(f) = \begin{cases} \bigcup f(S_x) \cup \{x\} & \text{if } \prec \text{ is a simple order on } \bigcup f(S_x) \cup \{x\} \\ \bigcup f(S_x) & \text{otherwise}\,. \end{cases}$$

Then by the general principle of recursive defamation (Exercise WO.1) there is a unique function $h : A \to \mathcal{P}(A)$ such that $h(\alpha) = \rho(h \restriction S_\alpha)$ for all $\alpha \in A$.

First we show that, for $\alpha, \beta \in A$ where $\alpha < \beta$, we have $h(\alpha) \subset h(\beta)$. So consider any $x \in h(\alpha) = \rho(h \restriction S_\alpha)$, and hence either $x \in \bigcup h(S_\alpha) \cup \{\alpha\}$ or $x \in \bigcup h(S_\alpha)$. Either way obviously $x \in \bigcup h(S_\alpha)$ so that there is a set $X \in h(S_\alpha)$ where $x \in X$. Then there is a $\gamma \in S_\alpha$ where $X = h(\gamma)$. Since we have $\alpha < \beta$, clearly also $\gamma \in S_\beta$ and hence $X \in h(S_\beta)$. Then also clearly both $x \in \bigcup h(S_\beta) \cup \{\beta\}$ and $x \in \bigcup h(S_\beta)$ so that for sure $x \in \rho(h \restriction S_\beta) = h(\beta)$. Since x was arbitrary this shows that $h(\alpha) \subset h(\beta)$ as desired.

Next we show by transfinite induction that $h(\alpha)$ is simply ordered by $\prec$ for every $\alpha \in A$. So consider $\alpha \in A$ and suppose that $h(\beta)$ is simply ordered by $\prec$ for every $\beta < \alpha$. If $\bigcup h(S_\alpha) \cup \{\alpha\}$ is simply ordered by $\prec$ then clearly $h(\alpha)$ is since then $h(\alpha) = \rho(h \restriction S_\alpha) = \bigcup h(S_\alpha) \cup \{\alpha\}$. So suppose that this is not the case so that $h(\alpha) = \rho(h \restriction S_\alpha) = \bigcup h(S_\alpha)$. Consider then any $x, y \in h(\alpha) = \bigcup h(S_\alpha)$ where $x \neq y$ so that there are X and Y in $h(S_\alpha)$ where $x \in X$ and $y \in Y$. Then there is a β and γ in S_α where $X = h(\beta)$ and $Y = h(\gamma)$. If $\beta = \gamma$ then x and y are both in $X = h(\beta) = h(\gamma) = Y$, which is simply ordered by the induction hypothesis so that x and y are comparable in $\prec$. If $\beta < \gamma$ then by what was shown above we have that $x \in X = h(\beta) \subset h(\gamma)$ so that x and y are both in $h(\gamma)$, which is simply ordered by the induction hypothesis so that again x and y are comparable. A similar argument shows that x and y are both in $h(\beta)$ and thus are comparable when $\beta > \gamma$. This completes the induction since x and y are comparable in all cases so that $h(\alpha)$ is always simply ordered.

We then claim that the set $B = \bigcup_{\alpha \in A} h(\alpha)$ is a maximal simply ordered (by $\prec$) subset of A, which of course shows the maximum principle. First, it is obviously a subset of A since each $h(\alpha) \in \mathcal{P}(A)$ so and so is a subset of A. To show that that B is simply ordered by $\prec$, consider x and y in B where $x \neq y$ so that there is an α and β in A where $x \in h(\alpha)$ and $y \in h(\beta)$. Without loss of generality we can assume that $\alpha < \beta$ so that $h(\alpha) \subset h(\beta)$ by what was shown below. Then both x and y are in $h(\beta)$, which is simply ordered by what was shown above. Hence x and y are comparable in $\prec$ so that B is simply ordered.

To show that B is maximal, suppose that $B \subsetneq Z$ and $Z \subset A$ is simply ordered by $\prec$. Then there is a $z \in Z$ where $z \notin B$. Now let $x \in \bigcup h(S_z)$ so that there is an $X \in h(S_z)$ where $x \in X$. Then there is an $\alpha \in S_z$ where $x \in X = h(\alpha)$. Hence clearly $x \in B$ so that also $x \in Z$ and so x and z are comparable in $\prec$ since Z is simply ordered. Since x was arbitrary this shows that the set $\bigcup h(S_z) \cup \{z\}$ is simply ordered so that $h(z) = \rho(h \restriction S_z) = \bigcup h(S_z) \cup \{z\}$. However, then we have that $z \in h(z)$ so that $z \in B$, which is a contradiction. So it must be that there is no such set Z and hence B is maximal. $\square$

Exercise WO.7

Use Exercises 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 $x \in T$,

$$x = c(X - S_x(T)) \,,$$

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 these two ordered sets are the same, or one equals a section of the other. [Hint: Switching indices if necessary, we can assume that $h : T_1 \to T_2$ is order preserving and $h(T_1)$ equals either T_2 or a section of T_2. Use Exercise 2 to show that $h(x) = x$ for all x.]

(b) If $(T, <)$ is a tower in X and $T \neq X$, show that 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 towers in X. Let

$$T = \bigcup_{k \in K} T_k \qquad \text{and} \qquad <= \bigcup_{k \in K} (<_k) \,.$$

Show that $(T, <)$ is a tower in X. Conclude that $T = X$.

Solution:

(a)

Proof. Since $(T_1, <_1)$ and $(T_2, <_2)$ are both well-ordered sets, it follows from Exercise WO.4a that either they have the same order type, T_1 has the same order type as a section of T_2, or vice-versa. We can assume that either they have the same order type of T_1 has the same order type as a section of T_2 since, in the third case, we can just swap the roles of T_1 and T_2. Thus there is an order preserving function $h : T_1 \to T_2$ whose image is either all of T_2 or a section of T_2. Given this, it was shown in the proof of Exercise WO.2a that $h(S_x(T_1)) = S_{h(x)}(T_2)$ for all $x \in T_1$.

We show that $h(x) = x$ for all $x \in T_1$ by transfinite induction. So suppose that $h(y) = y$ for all $y < x$, i.e. for all $y \in S_x(T_1)$ so that clearly $h(S_x(T_1)) = S_x(T_1)$. Then, since both T_1 and T_2 are towers in X and $h(x) \in T_2$, we have

$$h(x) = c(X - S_{h(x)}(T_2)) = c(X - h(S_x(T_1))) = c(X - S_x(T_1)) = x \,.$$

This completes the induction. Since $h(x) = x$ for all $x \in T_1$ and h preserves order, it follows that T_1 is equal to T_2 or a section of T_2 as desired. □

(b)

Proof. Since $T \neq X$, it follows that $X - T$ is nonempty. So let $a = c(X - T)$, $T' = T \cup \{a\}$, and $<' = < \cup \{(x, a) \mid x \in T\}$. Then clearly a is the largest element of T' and an upper bound of T so that $T = S_a(T')$, and hence $a = c(X - T) = c(X - S_a(T'))$. Since T is a tower, it then follows that T' is also a tower in X and that T is a section of T' as desired. □

(c)

Proof. First we need to show that $<$ is even a well-ordering of T as this is not obvious. To show that it is a simple order, consider $x, y \in T$ where $x \neq y$. It follows that $x \in T_k$ and $y \in T_l$ for some $k, l \in K$ by the definition of T. Since T_k and T_l are both towers in X, it follows from part (a) that they are equal or one is a section of the other. So, without loss of generality, we can assume that $T_k \subset T_l$ and also $<_k \subset <_l$. It then follows that both x and y are in T_l so that either $x <_l y$ or $y <_l x$ since $x \neq y$ and $<_l$ is a simple order. Then clearly $x < y$ or $y < x$ from the definition of $<$. This shows that $<$ has the comparability property.

Now suppose that there is an $x \in T$ where $x < x$. Then there is a $k \in K$ where $x <_k x$, which violates the fact that $<_k$ is a simple order. Hence it must be that $<$ is nonreflexive.

Lastly consider $x, y, z \in T$ where $x < y$ and $y < z$. Then there $k, l \in K$ where $x <_k y$ and $y <_l z$ and it must then be that $x, y \in T_k$ and $y, z \in T_l$. Again, since these are both towers, they are either equal or one is a section of the other by part (a). So we can assume that $T_k \subset T_l$ and $<_k \subset <_l$ without loss of generality so that we have $x, y, z \in T_l$ and $x <_l y <_l z$. From this clearly $x <_l z$ since it is a simple order and therefore transitive. Hence we have $x < y$, which of course shows that $<$ is transitive as well. This all shows that $<$ is indeed a simple order by definition.

To show that $<$ is a well-ordering, consider any nonempty subset Y of T. Then there is a $b \in Y$ so that also $b \in T$. It follows that there is a $k \in K$ such that $b \in T_k$, and also that $Y \cap T_k$ is a nonempty subset of T_k. It then follows that $Y \cap T_k$ has a smallest element a since T_k is well-ordered by $<_k$. We claim that in fact a is the smallest element of all of Y. To see this, consider any other $x \in Y$ so that also $x \in T$. Hence there is an $l \in K$ where $x \in T_l$. Now, since both T_k and T_l are towers in X, it follows from part (a) that they are equal or one is a section of the other.

Case: T_k and T_l are equal. Then both a and x are in T_k and so both in $Y \cap T_k$. Then $a \leq_k x$ since a is the smallest element of $Y \cap T_k$.

Case: T_k is a section of T_l. Then, if $a \in T_k$ then the argument in the previous case shows that $a \leq_k x$. On the other hand, if $a \notin T_k$, then it has to be that that $x <_l a$ since $x \in T_k$ and T_k is a section of T_l.

Case: T_l is a section of T_k. Then $T_l \subset T_k$ so that both a and x are in T_k and thus in $Y \cap T_k$. Hence again $a \leq_k x$ since a is the smallest element of $Y \cap T_k$.

In all cases $a \leq_m x$ for some $m \in K$ and hence $a \leq x$. Since x was an arbitrary element of Y, this shows that a is in fact the smallest element of Y. Since Y was an arbitrary nonempty subset of T, this shows that T is well-ordered by $<$.

Next we digress for a moment to show, for any $k \in K$ and $x \in T_k$, that $S_x(T_k) = S_x(T)$. So consider such k and x and suppose that $y \in S_x(T_k)$ so that $y <_k x$. Then clearly also $y < x$ by the definition of $<$ and hence $y \in S_x(T)$. This shows that $S_x(T_k) \subset S_x(T)$ since y was arbitrary. Now suppose $y \in S_x(T)$ so that $y < x$. Then there is an $l \in K$ where $y <_l x$. Hence $x, y \in T_l$, and by part (a) either T_l and T_k are equal, or one is a section of the other. If they are equal or T_l is a section of T_k then clearly we have $<_l \subset <_k$ so that $y <_k x$. If T_k is a section of T_l then, since $y <_l x$ and $x \in T_k$, it has to be that also $y \in T_k$ since T_k is a section of T_l. Hence it must be that $y <_k x$. Since this is true in all cases it follows that $y \in S_x(T_k)$, which shows that $S_x(T) \subset S_x(T_k)$. This completes the proof that $S_x(T_k) = S_x(T)$.

With this having been shown, we can easily show that T is a tower in X. For any $x \in T$ there is a $k \in K$ where $x \in T_k$. Since T_k is a tower in X we have

$$x = c(X - S_x(T_k)) = c(X - S_x(T))$$

by what was just shown. Thus suffices to show that T is a tower in X.

Lastly, we claim that $T = X$. To see this, suppose that it is not the case so that by part (b) there is a tower S in X such that T is a section of S. From this we have that $T = S_a(S)$ for some $a \in S$ and that of course $a \notin T$. However, since S is a tower and $\{(T_k, <_k) \mid k \in K\}$ is the collection of *all* towers in X, it follows that there must be a $k \in K$ such that $S = T_k$. Then we have that $a \in S = T_k$ so that of course $a \in T$ by definition, which is a contradiction. So it must be that in fact $T = X$ as desired.

This of course shows that $<$ is a well-ordering of $X = T$ so that the choice axiom implies the well-ordering theorem since X is an arbitrary set. In contrast to the previous proof, it is easy to prove that the well-ordering theorem implies the choice axiom. For a collection of nonempty sets $\mathcal{B}$ define $X = \bigcup \mathcal{B}$. Then X can be well-ordered by the well-ordering theorem. Then we simply define a choice function c on $\mathcal{B}$ in the following way: any $B \in \mathcal{B}$ is clearly a nonempty subset of X and so has a smallest element a since X is well ordered. So simply set $c(B) = a$, from which it is clear that $c(B) \in B$ and so c is a valid choice function. $\square$

Exercise WO.8

Using Exercises 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 $(A, <) \sim (A', <')$ if $(A, <)$ and $(A', <')$ have the same order type. It is trivial to show that this is an equivalence relation. Let $[(A, <)]$ denote the equivalence class of $(A, <)$; let E denote the collection of these equivalence classes. Define

$$[(A, <)] \ll [(A', <')]$$

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 $[(\varnothing, \varnothing)]$ 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. [Hint: Define a map $f : A \to E$ by setting $f(x) = [(S_x(A), \text{restriction } <)]$ for each $x \in A$.]

(c) Conclude that E is well-ordered by $\ll$.

(d) Show that E is uncountable. [Hint: If $h : E \to \mathbb{Z}_+$ is a bijection, then h gives rise to a well-ordering of $\mathbb{Z}_+$.]

Solution:

(a)

Proof. First, to show that $\ll$ is well defined, suppose that $[(A, <)] \ll [(A', <')]$ and that $(B, \prec) \in [(A, <)]$ and $(B', \prec') \in [(A', <')]$. Then $(A, <)$ has the same order type as a section of $(A', <')$ so that there is an order-preserving map h from A onto a section of A'. We also then have that $(B, \prec)$ has the same order type as $(A, <)$ since they are in the same equivalence class. Thus there is an order-preserving bijection $f : B \to A$. Likewise there is an order-preserving bijection from $g : B' \to A'$. It is then trivial to show that $g^{-1} \circ h \circ f$ is bijection from B onto a section of B' that preserves order. Hence $(B, \prec)$ has the same order type as a section of $(B', \prec')$. Since $(B, \prec)$ and $(B', \prec')$ were arbitrary elements in their respective equivalence classes, this shows that $\ll$ is well defined such that it does not matter which representatives we use from the equivalence classes.

Now consider any equivalence class $[(A, <)]$ in E. Then clearly it cannot be that $[(A, <)] \ll [(A, <)]$, since this would mean that A has the same order type as a section of itself, which would contradict what was shown in Exercise WO.2b. Thus $\ll$ is nonreflexive.

Next consider two distinct equivalence classes $[(A, <)]$ and $[(A', <')]$. Then it cannot be that $(A, <)$ and $(A', <')$ have the same order type, for then they would be the same equivalence class. Then, by

Exercise WO.4a, it must be that either $(A, <)$ has the same order type as a section of $(A', <')$ or vice-versa. Clearly then, in the former case $[(A, <)] \ll [(A', <')]$, and in the latter case $[(A', <')] \ll [(A, <)]$. This shows that $\ll$ has the comparability property.

Lastly, suppose that $[(A, <)] \ll [(A', <')]$ and $[(A', <')] \ll [(A'', <'')]$. Then $(A, <)$ has the same order type as section of $(A', <')$ so that there is an order-preserving bijection f from A onto a section of A' Likewise there is an order-preserving bijection g from A' onto a section of A''. It is then trivial to show that $f \circ g$ is an order-preserving bijection from A onto a section of A''. It then clearly follows that $[(A, <)] \ll [(A'', <'')]$, which shows that $\ll$ is transitive.

Hence we have shown that $\ll$ satisfies all the requirements of a simple order. □

(b)

Proof. Following the hint, define the map $f : A \to E$ by setting

$$f(x) = [(S_x(A), \text{restriction } <)]$$

for any $x \in A$, noting that clearly $S_x(A)$ is well-ordered by the restricted $<$ so that the equivalence class is valid and in E.

Consider any x and y in A where $x < y$. Then clearly $x \in S_y(A)$ but $x \notin S_x(A)$ (since it is not true that $x < x$) so that $S_x(A)$ and $S_y(A)$ are distinct sets. We also clearly have that $S_x(A) = S_x(S_y(A))$ so that $S_x(A)$ has the same order type (the identity function is the required order-preserving map) as a section of $S_y(A)$. Hence

$$f(x) = [(S_x(A), \text{restriction } <)] \ll [(S_y(A), \text{restriction } <)] = f(y)$$

so that f preserves order since x and y were arbitrary.

Now we show that f is onto $S_\alpha(E)$. So consider any equivalence class $[(B, \prec)]$ in $S_\alpha(E)$ and hence

$$[(B, \prec)] \ll \alpha = [(A, <)]$$

so that by definition $(B, \prec)$ has the same order type as some section $S_x(A)$. Hence $[(B, \prec)]$ and $[(S_x(A), \text{restriction } <)]$ are the same equivalence class! Therefore

$$f(x) = [(S_x(A), \text{restriction } <)] = [(B, \prec)] \, ,$$

which of course shows the desired property since $[(B, \prec)]$ was arbitrary.

This shows that f is an order-preserving map from A onto $S_\alpha(E)$ so that they have the same order type. □

(c)

Proof. Consider any nonempty subset $D \subset E$. Thus there is an $\alpha = [(A, <)] \in D$. If α is the smallest element of D then we are done, so assume that this is not the case so that there is a $\beta \in D$ where $\beta \ll \alpha$. Now, it was shown in part (b) that $(A, <)$ has the same order type as the section $S_\alpha(E)$ so that this section must be well-ordered since A is. Also we have that $\beta \in S_\alpha(E)$ since $\beta \ll \alpha$. Thus $\beta \in D \cap S_\alpha(E)$ so that $D \cap S_\alpha(E)$ is a nonempty subset of $S_\alpha(E)$ so has a smallest element γ since $S_\alpha(E)$ is well-ordered. In particular, we of course have that $\gamma \lll \beta$, where we use $\lll$ to denote $\ll$ or equal to.

We claim that γ must be the smallest element of D. If not, then there is a $\delta \in D$ where $\delta \ll \gamma$. Of course we also then have that $\delta \ll \gamma \lll \beta \ll \alpha$ and hence $\delta \in S_\alpha(E)$. Therefore $\delta \in D \cap S_\alpha(E)$, but since $\delta \ll \gamma$ this contradicts the definition of γ as the smallest element of $D \cap S_\alpha(E)$. So it must be that in fact γ is the smallest element of D, which shows that E is well-ordered by $\ll$ since D was an arbitrary subset. □

(d)

Proof. Following the hint, suppose that E is countable so that there is a bijection $h : E \to \mathbb{Z}_+$. This of course gives rise to a well-ordering $<$ of $h(E) = \mathbb{Z}_+$ by simply ordering its elements according to its bijection with E, which was shown to be well-ordered in part (c). Then we have that $(\mathbb{Z}_+, <)$ is an element of $\mathcal{A}$ since $\mathbb{Z}_+$ is a subset of itself. Thus the equivalence class $\alpha = [(\mathbb{Z}_+, <)]$ is an element of E. But we know from part (b) that $(\mathbb{Z}_+, <)$ then has the same order type as the section $S_\alpha(E)$. Since we also know that $(\mathbb{Z}_+, <)$ has the same order type as E itself, it follows that E has the same order type as its section $S_\alpha(E)$. This was shown not to be possible in Exercise WO.2b so that a contradiction has been reached. So it must be that in fact E is uncountable as desired! □

Chapter 2 Topological Spaces and Continuous Functions

§13 Basis for a Topology

Exercise 13.1

Let X be a topological space; let A be a subset of X. Suppose that for each $x \in A$ there is an open set U containing x such that $U \subset A$. Show that A is open in X.

Solution:

Proof. For each $x \in A$ we can choose an open set U_x containing x such that $U_x \subset A$. We then claim that $\bigcup_{x \in A} U_x = A$. So first consider any $y \in \bigcup_{x \in A} U_x$ so that there is an $x \in A$ such that $y \in U_x$. Then clearly also $y \in A$ since $U_x \subset A$. Hence $\bigcup_{x \in A} U_x \subset A$ since y was arbitrary. Now consider $y \in A$ so that clearly $y \in U_y$. Then obviously $y \in \bigcup_{x \in A} U_x$ so that $A \subset \bigcup_{x \in A} U_x$ since y was arbitrary. Thus we have shown that $\bigcup_{x \in A} U_x = A$, and since each U_x is open, it follows from the definition of a topology that the union $\bigcup_{x \in A} U_x = A$ is open as well. □

Exercise 13.2

Consider the nine topologies on the set $X = \{a, b, c\}$ indicated in Example 1 of §12. Compare them; that is, for each pair of topologies, determine whether they are comparable, and if so, which is finer.

Solution:

We label each of the topologies in Figure 12.1 with an ordered pair (i, j) where $1 \le i, j \le 3$, i is the row, j is the column, and $(1, 1)$ is the upper left corner. The following matrix lists which of each pair is *finer*, or "Inc" if they are incomparable.

	$(1,1)$	$(1,2)$	$(1,3)$	$(2,1)$	$(2,2)$	$(2,3)$	$(3,1)$	$(3,2)$	$(3,3)$
$(1,1)$	$=$	$(1,2)$	$(1,3)$	$(2,1)$	$(2,2)$	$(2,3)$	$(3,1)$	$(3,2)$	$(3,3)$
$(1,2)$		$=$	Inc	Inc	Inc	Inc	$(1,2)$	$(3,2)$	$(3,3)$
$(1,3)$			$=$	$(1,3)$	Inc	$(2,3)$	$(1,3)$	Inc	$(3,3)$
$(2,1)$				$=$	Inc	$(2,3)$	Inc	$(3,2)$	$(3,3)$
$(2,2)$					$=$	Inc	Inc	Inc	$(3,3)$
$(2,3)$						$=$	$(2,3)$	Inc	$(3,3)$
$(3,1)$							$=$	$(3,2)$	$(3,3)$
$(3,2)$								$=$	$(3,3)$
$(3,3)$									$=$

We know that $\subsetneq$ forms a strict partial order on these topologies. So we can also list all the maximal simply ordered subsets, each in order:

$$(1,1) \subsetneq (2,2) \subsetneq (3,3)$$
$$(1,1) \subsetneq (3,1) \subsetneq (1,2) \subsetneq (3,2) \subsetneq (3,3)$$
$$(1,1) \subsetneq (3,1) \subsetneq (1,3) \subsetneq (2,3) \subsetneq (3,3)$$
$$(1,1) \subsetneq (2,1) \subsetneq (1,3) \subsetneq (2,3) \subsetneq (3,3)$$
$$(1,1) \subsetneq (2,1) \subsetneq (3,2) \subsetneq (3,3)$$

Exercise 13.3

Show that the collection $\mathcal{T}_\alpha$ given in Example 4 of §12 is a topology on X. Is the collection

$$\mathcal{T}_\infty = \{U \mid X - U \text{ is infinite or empty or all of } X\}$$

a topology on X?

Solution:

Recall that $\mathcal{T}_\alpha$ from Example 12.4 is the set of all subsets U of X such that $X - U$ either is countable or is all of X. First we show that $\mathcal{T}_\alpha$ is a topology on X.

Proof. First, clearly $\varnothing \in \mathcal{T}_\alpha$ since $X - \varnothing = X$ is all of X. Also $X \in \mathcal{T}_\alpha$ since $X - X = \varnothing$ is countable. Now suppose that $\mathcal{A}$ is a subcollection of $\mathcal{T}_\alpha$ so that $X - U$ is countable (or all of X) for every $U \in \mathcal{A}$. Then we have that

$$X - \bigcup \mathcal{A} = X - \bigcup_{A \in \mathcal{A}} A = \bigcap_{A \in \mathcal{A}} (X - A)$$

is countable (or all of X) since every $X - A$ is countable (or all of X). Therefore $\bigcup \mathcal{A} \in \mathcal{T}_\alpha$ by definition.

Now suppose that $U_1, \ldots U_n$ are nonempty elements of $\mathcal{T}_\alpha$ so that $X - U_i$ is a countable subset of X or X itself for each $i \in \{1, \ldots, n\}$. Then we have

$$X - \bigcap_{i=1}^{n} U_i = \bigcup_{i=1}^{n} (X - U_i)$$

is a finite union of sets that are either countable subsets of X, or X itself. It then follows that the union is countable or X itself so that $\bigcap_{i=1}^{n} U_i \in \mathcal{T}_\alpha$ by definition. This completes the proof that $\mathcal{T}_\alpha$ is a topology on X. $\square$

Now we claim that the collection $\mathcal{T}_\infty$ as defined above is not always a topology on X.

Proof. As a counterexample, let $X = \mathbb{Z}_+$ and suppose that $\mathcal{T}_\infty$ is a topology on X. Clearly if U is a finite subset of X, then $X - U$ is infinite since X is infinite so that U is open. Now consider the subcollection

$$\mathcal{A} = \{\{i\} \mid i \in \mathbb{Z}_+ \text{ and } i > 1\} = \{\{2\}, \{3\}, \ldots\}\ .$$

Then clearly we have that $\bigcup \mathcal{A} = \{2, 3, \ldots\}$ so that $X - \bigcup \mathcal{A} = \{1\}$ is neither infinite, empty, nor all of X. Therefore $\bigcup \mathcal{A}$ cannot be open, which violates property (2) of a topology. So it must be that $\mathcal{T}_\infty$ is not a topology, which of course contradicts our supposition that it is! $\square$

Exercise 13.4

(a) If $\{\mathcal{T}_\alpha\}$ is a family of topologies on X, show that $\bigcap \mathcal{T}_\alpha$ is a topology on X. Is $\bigcup \mathcal{T}_\alpha$ a topology on X?

(b) Let $\{\mathcal{T}_\alpha\}$ be family of topologies on X. Show that there is a unique smallest topology on X containing all the collections $\mathcal{T}_\alpha$, and a unique largest topology contained in all $\mathcal{T}_\alpha$.

(c) If $X = \{a, b, c\}$, let

$$\mathcal{T}_1 = \{\varnothing, X, \{a\}, \{a, b\}\} \qquad \text{and} \qquad \mathcal{T}_2 = \{\varnothing, X, \{a\}, \{b, c\}\} .$$

Find the smallest topology containing $\mathcal{T}_1$ and $\mathcal{T}_2$, and the largest topology contained in $\mathcal{T}_1$ and $\mathcal{T}_2$.

Solution:

(a) First we show that $\bigcap \mathcal{T}_\alpha$ is a topology on X.

Proof. First, clearly since $\varnothing$ and X are in every $\mathcal{T}_\alpha$ since they are topologies, they are both in $\bigcap \mathcal{T}_\alpha$ so that property (1). Now suppose that $\mathcal{A}$ is a subcollection of $\bigcap \mathcal{T}_\alpha$. Consider any $\mathcal{T}_\beta$ and any $A \in \mathcal{A}$. Then A is also in $\bigcap \mathcal{T}_\alpha$ since $\mathcal{A} \subset \bigcap \mathcal{T}_\alpha$. It then follows that A is in our specific $\mathcal{T}_\beta$. Since A was arbitrary it follows that $\mathcal{A}$ is a subcollection of $\mathcal{T}_\beta$ so that $\bigcup \mathcal{A} \in \mathcal{T}_\beta$ also since $\mathcal{T}_\beta$ is a topology. Since $\mathcal{T}_\beta$ was also arbitrary it follows that $\bigcup \mathcal{A} \in \bigcap \mathcal{T}_\alpha$. Lastly, since the subcollection $\mathcal{A}$ was arbitrary, this shows property (2) for $\bigcap \mathcal{T}_\alpha$.

Finally, suppose that $U_1, \ldots, U_n$ are sets in $\bigcap \mathcal{T}_\alpha$. Consider any $\mathcal{T}_\beta$ so that clearly then $U_i \in \mathcal{T}_\beta$ for every $i \in \{1, \ldots, n\}$. It then follows that $\bigcap_{i=1}^n U_i \in \mathcal{T}_\beta$ since $\mathcal{T}_\beta$ is a topology. Since $\mathcal{T}_\beta$ was arbitrary, this shows that $\bigcap_{i=1}^n U_i \in \bigcap \mathcal{T}_\alpha$, which shows property (3) for $\bigcap \mathcal{T}_\alpha$. This completes the proof that $\bigcap \mathcal{T}_\alpha$ is a topology on X since all three properties have been shown. $\square$

Now we claim that $\bigcup \mathcal{T}_\alpha$ is *not* generally a topology.

Proof. As a counterexample consider the set $X = \{a, b, c\}$, the topologies $\mathcal{T}_1 = \{\varnothing, X, \{a\}\}$ and $\mathcal{T}_2 = \{\varnothing, X, \{b\}\}$, and the collection of topologies $\mathcal{C} = \{\mathcal{T}_1, \mathcal{T}_2\}$. Then we clearly have that $\bigcup \mathcal{C} = \mathcal{T}_1 \cup \mathcal{T}_2 = \{\varnothing, X, \{a\}, \{b\}\}$, which is not a topology since $\mathcal{A} = \{\{a\}, \{b\}\}$ is a subcollection of $\bigcup \mathcal{C}$ but $\bigcup \mathcal{A} = \{a, b\}$ is not in $\bigcup \mathcal{C}$. $\square$

(b) First we show that there is a unique smallest topology that contains each $\mathcal{T}_\alpha$.

Proof. It was proven in part (a) that $\bigcup \mathcal{T}_\alpha$ is not necessarily a topology. However, it is clearly always a subbasis for a topology since clearly $X \in \bigcup \mathcal{T}_\alpha$ since it is in each $\mathcal{T}_\alpha$ since they are topologies. Hence obviously then $\bigcup (\bigcup \mathcal{T}_\alpha) = X$ so that $\bigcup \mathcal{T}_\alpha$ is a subbasis by definition. Then let $\mathcal{T}_s$ be the topology generated by the subbasis $\bigcup \mathcal{T}_\alpha$. We claim that $\mathcal{T}_s$ is then the smallest topology that contains all the $\mathcal{T}_\alpha$ as subsets.

First, from the proof following the definition of a subbasis, we know that the set $\mathcal{B}$ of finite intersections of elements of $\bigcup \mathcal{T}_\alpha$ is a basis for the topology $\mathcal{T}_s$, and that $\mathcal{T}_s$ is the set of all unions of subcollections of $\mathcal{B}$.

We first show that every $\mathcal{T}_\alpha$ is indeed contained as a subset of $\mathcal{T}_s$. So consider any specific $\mathcal{T}_\beta$ and any $U \in \mathcal{T}_\beta$. Then clearly $U \in \bigcup \mathcal{T}_\alpha$ so that $U \in \mathcal{B}$ since $U = \bigcap \{U\}$ is a finite intersection of elements of $\bigcup \mathcal{T}_\alpha$. It then follows that $U \in \mathcal{T}_s$ since $U = \bigcup \{U\}$ is the union of a subcollection of $\mathcal{B}$. Since U was arbitrary, this shows that $\mathcal{T}_\beta \subset \mathcal{T}_s$, which shows the result since $\mathcal{T}_\beta$ was arbitrary.

Now we show that $\mathcal{T}_s$ is the smallest such topology as ordered by $\subsetneq$. So suppose that $\mathcal{T}$ is a topology that contains every $\mathcal{T}_\alpha$ as a subset. Consider any $U \in \mathcal{T}_s$ so that $U = \bigcup \mathcal{C}$ for some subcollection

$\mathcal{C} \subset \mathcal{B}$. Now consider any $Y \in \mathcal{C}$ so that also $Y \in \mathcal{B}$. Then $Y = \bigcap_{i=1}^{n} Y_i$ where each $Y_i \in \bigcup \mathcal{T}_\alpha$. Then each Y_i is in some $\mathcal{T}_\beta \subset \mathcal{T}$ so that also $Y_i \in \mathcal{T}$. Since $\mathcal{T}$ is a topology, it follows that the finite intersection $\bigcap_{i=1}^{n} Y_i = Y$ is also in $\mathcal{T}$. Since Y was arbitrary, this shows that $\mathcal{C} \subset \mathcal{T}$ so that $\mathcal{C}$ is a subcollection of $\mathcal{T}$. It then follows that $\bigcup \mathcal{C} = U$ is also in $\mathcal{T}$ since $\mathcal{T}$ is a topology. Since U was arbitrary, we have that $\mathcal{T}_s \subset \mathcal{T}$, which shows that $\mathcal{T}_s$ is the smallest topology since $\mathcal{T}$ was arbitrary.

It is easy to see that $\mathcal{T}_s$ is unique since, if both $\mathcal{T}_1$ and $\mathcal{T}_s$ are the smallest topologies that contain each $\mathcal{T}_\alpha$ as subsets, then we would have that both $\mathcal{T}_1 \subset \mathcal{T}_2$ and $\mathcal{T}_2 \subset \mathcal{T}_1$ so that $\mathcal{T}_1 = \mathcal{T}_2$. Really this follows from the more general fact that smallest elements in any order are always unique. □

Next we show that there is a unique largest topology that is contained in each $\mathcal{T}_\alpha$.

Proof. It was shown in part (a) that $\mathcal{T}_l = \bigcap \mathcal{T}_\alpha$ is a topology on X. We claim that in fact this is the unique largest topology contained in all $\mathcal{T}_\alpha$. First, clearly $\mathcal{T}_l = \bigcap \mathcal{T}_\alpha$ is contained in each $\mathcal{T}_\alpha$ since the intersection of a collection of sets is always a subset of every set in the collection. Now suppose that $\mathcal{T}$ is a topology that is contained in every $\mathcal{T}_\alpha$, i.e. $\mathcal{T} \subset \mathcal{T}_\alpha$ for every $\mathcal{T}_\alpha$. Then clearly for any $U \in \mathcal{T}$ we have that $U \in \mathcal{T}_\alpha$ for every $\mathcal{T}_\alpha$ so that $U \in \bigcap \mathcal{T}_\alpha = \mathcal{T}_l$. Thus $\mathcal{T} \subset \mathcal{T}_l$ since U was arbitrary. This shows that $\mathcal{T}_l$ is the largest such topology since $\mathcal{T}$ was arbitrary.

Clearly also $\mathcal{T}_l$ is unique since, if $\mathcal{T}_1$ and $\mathcal{T}_2$ are two such largest topologies that are contained in every $\mathcal{T}_\alpha$. Then we would have $\mathcal{T}_1 \subset \mathcal{T}_2$ and $\mathcal{T}_2 \subset \mathcal{T}_1$ so that $\mathcal{T}_1 = \mathcal{T}_2$. This also follows from the fact that the largest element in any ordered set (or collection of sets in this case) is unique. □

(c) Note that the proofs in part (b) are constructive so that we can construct these topologies as done in the proof. For the smallest topology containing $\mathcal{T}_1$ and $\mathcal{T}_2$ we have that

$$\bigcup \{\mathcal{T}_1, \mathcal{T}_2\} = \mathcal{T}_1 \cup \mathcal{T}_2 = \{\varnothing, X, \{a\}, \{a,b\}, \{b,c\}\}$$

is a subbasis for the smallest topology $\mathcal{T}_s$. Then the collection of all finite intersections of elements of this set is a basis for $\mathcal{T}_s$:

$$\mathcal{B} = \{\varnothing, X, \{a\}, \{b\}, \{a,b\}, \{b,c\}\} .$$

Then the topology $\mathcal{T}_s$ is the set of all unions of subcollections of $\mathcal{B}$:

$$\mathcal{T}_s = \{\varnothing, X, \{a\}, \{b\}, \{a,b\}, \{b,c\}\} = \mathcal{B}$$

so that evidently the basis and the topology are the same set here!

For the largest topology contained in $\mathcal{T}_1$ and $\mathcal{T}_2$ we have simply

$$\mathcal{T}_l = \bigcap \{\mathcal{T}_1, \mathcal{T}_2\} = \mathcal{T}_1 \cap \mathcal{T}_2 = \{\varnothing, X, \{a\}\} .$$

Exercise 13.5

Show that if $\mathcal{A}$ is a basis for a topology on X, then the topology generated by $\mathcal{A}$ equals the intersection of all topologies on X that contain $\mathcal{A}$. Prove the same if $\mathcal{A}$ is a subbasis.

Solution:

Suppose that $\mathcal{T}$ is the topology generated by basis $\mathcal{A}$, and $\mathcal{C}$ is the collection of topologies on X that contain $\mathcal{A}$ as a subset.

First we show that $\mathcal{T} = \bigcap \mathcal{C}$.

Proof. Consider $U \in \mathcal{T}$ and any $\mathcal{T}_c \in \mathcal{C}$ so that $\mathcal{A} \subset \mathcal{T}_c$. Then, since $\mathcal{A}$ generates $\mathcal{T}$, it follows from Lemma 13.1 that U is the union of elements of $\mathcal{A}$. Clearly then each of these elements of $\mathcal{A}$ is in $\mathcal{T}_c$ since $\mathcal{A} \subset \mathcal{T}_c$ so that their union is as well since $\mathcal{T}_c$ is a topology. Hence $U \in \mathcal{T}_c$ so that $\mathcal{T} \subset \mathcal{T}_c$ since U was arbitrary. Hence $\mathcal{T}$ is contained in all elements of $\mathcal{C}$ so that $\mathcal{T} \subset \bigcap \mathcal{C}$. Also, clearly $\mathcal{T}$ is a topology that contains $\mathcal{A}$ so that $\mathcal{T} \in \mathcal{C}$. Clearly then $\bigcap \mathcal{C} \subset \mathcal{T}$ so that $\mathcal{T} = \bigcap \mathcal{C}$ as desired. □

Next we show the same thing but when $\mathcal{A}$ is a subbasis.

Proof. Let $\mathcal{B}$ be the set of all finite intersections of elements of $\mathcal{A}$, which we know is a basis for $\mathcal{T}$ by the proof after the definition of a subbasis. We show that $\mathcal{B} \subset \mathcal{T}_c$ for all $\mathcal{T}_c \in \mathcal{C}$. So consider any set $B \in \mathcal{B}$ so that B is the finite intersection of elements of $\mathcal{A}$. Also consider any $\mathcal{T}_c \in \mathcal{C}$ so that each of these elements is in $\mathcal{T}_c$ since $\mathcal{A} \subset \mathcal{T}_c$. Since $\mathcal{T}_c$ is a topology, clearly the finite intersection of these elements, i.e. B, is in $\mathcal{T}_c$. Hence $\mathcal{B} \subset \mathcal{T}_c$ since B was arbitrary.

It then follows from what was shown before that $\mathcal{T} = \bigcap \mathcal{C}$ since $\mathcal{T}$ is the topology generated by the basis $\mathcal{B}$ and $\mathcal{B}$ is contained in each topology in $\mathcal{C}$. □

Exercise 13.6

Show that the topologies of $\mathbb{R}_l$ and $\mathbb{R}_K$ are not comparable.

Solution:

Proof. Let $\mathcal{T}_l$ and $\mathcal{T}_K$ be the topologies of $\mathbb{R}_l$ and $\mathbb{R}_K$, respectively. Also let $\mathcal{B}_l$ and $\mathcal{B}_K$ be the corresponding bases.

Consider $x = 0 \in \mathbb{R}$ and $B_l = [0,1)$, which clearly contains 0 and is a basis element of $\mathcal{B}_l$. Let B_K be any basis element of $\mathcal{B}_K$ that contains 0. Then B_K is either (a,b) or $(a,b) - K$ for some $a < b$. In either case it must be that $a < 0 < b$ so that clearly $a < a/2 < 0 < b$. Also $a/2 \notin K$ since $a/2 < 0$ so that we have $a/2 \in (a,b)$ and $a/2 \in (a,b) - K$. Clearly also $a/2 \notin [0,1)$ so that it cannot be that $B_K \subset B_l$. We have therefore shown that

$$\begin{gathered}
\exists x \in \mathbb{R} \exists B_l \in \mathcal{B}_l \left[x \in B_l \wedge \forall B_K \in \mathcal{B}_K \left(x \in B_K \Rightarrow B_K \not\subset B_l\right)\right] \\
\exists x \in \mathbb{R} \exists B_l \in \mathcal{B}_l \left[x \in B_l \wedge \forall B_K \in \mathcal{B}_K \left(x \notin B_K \vee B_K \not\subset B_l\right)\right] \\
\exists x \in \mathbb{R} \exists B_l \in \mathcal{B}_l \left[x \in B_l \wedge \neg \exists B_K \in \mathcal{B}_K \left(x \in B_K \wedge B_K \subset B_l\right)\right] \\
\exists x \in \mathbb{R} \exists B_l \in \mathcal{B}_l \neg \left[x \notin B_l \vee \exists B_K \in \mathcal{B}_K \left(x \in B_K \wedge B_K \subset B_l\right)\right] \\
\exists x \in \mathbb{R} \exists B_l \in \mathcal{B}_l \neg \left[x \in B_l \Rightarrow \exists B_K \in \mathcal{B}_K \left(x \in B_K \wedge B_K \subset B_l\right)\right] \\
\neg \forall x \in \mathbb{R} \forall B_l \in \mathcal{B}_l \left[x \in B_l \Rightarrow \exists B_K \in \mathcal{B}_K \left(x \in B_K \wedge B_K \subset B_l\right)\right] \\
\neg \forall x \in \mathbb{R} \forall B_l \in \mathcal{B}_l \left[x \in B_l \Rightarrow \exists B_K \in \mathcal{B}_K \left(x \in B_K \subset B_l\right)\right]
\end{gathered}$$

This shows by the negation of Lemma 13.3 that $\mathcal{T}_K$ is not finer than $\mathcal{T}_l$.

Now consider again $x = 0 \in \mathbb{R}$ and $B_K = (-1,1) - K$, which clearly contains 0 and is a basis element of $\mathcal{B}_K$. Let B_l be any basis element of $\mathcal{B}_l$ that contains 0 so that $B_l = [a,b)$ where $a \le 0 < b$. Clearly we have that $1/b > 0$ and there is an $n \in \mathbb{Z}_+$ where $n > 1/b$ since the positive integers have no upper bound. We then have

$$\begin{aligned}
&0 < 1/b < n \\
&0 < 1 < bn && \text{(since } b > 0\text{)} \\
&0 < 1/n < b && \text{(since } n > 1/b > 0\text{)}
\end{aligned}$$

so that $1/n \in [0, b) = B_l$. However, clearly $1/n \in K$ so that $1/n \notin (-1, 1) - K = B_K$. Hence it must be that $B_l \not\subset B_K$. This shows that $\mathcal{T}_l$ is not finer than $\mathcal{T}_K$ by the negation of Lemma 13.3 as before.

This completes the proof that $\mathcal{T}_K$ and $\mathcal{T}_l$ are not comparable. □

Exercise 13.7

Consider the following topologies on $\mathbb{R}$:

$$\begin{aligned}
\mathcal{T}_1 &= \text{the standard topology,} \\
\mathcal{T}_2 &= \text{the topology of } \mathbb{R}_K, \\
\mathcal{T}_3 &= \text{the finite compliment topology,} \\
\mathcal{T}_4 &= \text{the upper limit topology, having all sets } (a, b] \text{ as basis,} \\
\mathcal{T}_5 &= \text{the topology having all sets } (-\infty, a) = \{x \mid x < a\} \text{ as basis.}
\end{aligned}$$

Determine, for each of these topologies, which of the others it contains.

Solution:

We claim that $\mathcal{T}_3 \subsetneq \mathcal{T}_1 \subsetneq \mathcal{T}_2 \subsetneq \mathcal{T}_4$ and $\mathcal{T}_5 \subsetneq \mathcal{T}_1 \subsetneq \mathcal{T}_2 \subsetneq \mathcal{T}_4$ but that $\mathcal{T}_3$ and $\mathcal{T}_5$ are incomparable.

Let $\mathcal{B}_1$, $\mathcal{B}_2$, $\mathcal{B}_4$, and $\mathcal{B}_5$ be the given bases corresponding to the above topologies, noting that $\mathcal{T}_3$ is defined directly rather than generated from a basis.

First we show that $\mathcal{T}_3 \subsetneq \mathcal{T}_1$.

Proof. Consider any $U \in \mathcal{T}_3$ so that $\mathbb{R} - U$ is finite or $U = \mathbb{R}$. Clearly in the latter case $U \in \mathcal{T}_1$ since it is a topology. In the former case $\mathbb{R} - U$ is a finite set of real numbers so that its elements can be enumerated as $\{x_1, x_2, \ldots, x_n\}$ for some $n \in \mathbb{Z}_+$ where $x_1 < x_2 < \cdots < x_n$. Then clearly we have that

$$U = (-\infty, x_1) \cup \left[\bigcup_{k=1}^{n-1} (x_k, x_{k+1})\right] \cup (x_n, \infty) .$$

Each of these sets is an interval (a, b) or the union of such intervals. For example, the set $(-\infty, x_1)$ can be covered by the countable union of intervals

$$\bigcup_{k=1}^{\infty} (x_1 - k - 1, x_1 - k + 1)$$

and similarly for the interval (x_n, ∞). Hence the union U is an element of $\mathcal{T}_1$ by Lemma 13.1. Since U was arbitrary, this shows that $\mathcal{T}_3 \subset \mathcal{T}_1$.

Now, clearly the interval $(-1, 1)$ is in $\mathcal{T}_1$ since it is a basis element. However, we also have that $\mathbb{R} - (-1, 1) = (-\infty, -1] \cup [1, \infty)$ is neither finite nor all of $\mathbb{R}$. Hence $(-1, 1) \notin \mathcal{T}_3$. This shows that $\mathcal{T}_1$ cannot be a subset of $\mathcal{T}_3$ so that $\mathcal{T}_3 \subsetneq \mathcal{T}_1$ as desired. □

Next we show that $\mathcal{T}_5 \subsetneq \mathcal{T}_1$ also.

Proof. Consider any $x \in \mathbb{R}$ and any basis element $B_5 \in \mathcal{B}_5$ containing x. Then $B_5 = (-\infty, a)$ where $x < a$. Let $B_1 = (x - 1, a)$, which is a basis element in $\mathcal{B}_1$. Also clearly B_1 contains x and is a subset of B_5. This proves that $\mathcal{T}_5 \subset \mathcal{T}_1$ by Lemma 13.3.

Now consider $x = -1$ and basis element $B_1 = (-2, 0)$ in $\mathcal{B}_1$, noting that obviously $x \in B_1$, and hence $-2 < x < 0$. Let B_5 be any element of $\mathcal{B}_5$ containing x so that $B_5 = (-\infty, a)$ where $x < a$. Clearly then $-3 < -2 < x < a$ so that $-3 \in B_5$. However, since $-3 \notin (-2, 0) = B_1$, this shows that $B_5 \not\subset B_1$. This suffices to show that $\mathcal{T}_1 \not\subset \mathcal{T}_5$ by the negation of Lemma 13.3. Therefore $\mathcal{T}_5 \subsetneq \mathcal{T}_1$ as desired. □

Now we show that $\mathcal{T}_3$ and $\mathcal{T}_5$ are not comparable.

Proof. First consider the set $U = \mathbb{R} - \{0\}$ so that $U \in \mathcal{T}_3$ since $\mathbb{R} - U = \{0\}$ is obviously finite. Now suppose that $U \in \mathcal{T}_5$ as well. Then, since clearly $1 \in U$, there must be a basis element $B_5 \in \mathcal{B}_5$ where $1 \in B_5$ and $B_5 \subset U$ by the definition of a topological basis. Then $B_5 = (-\infty, a)$ where $1 < a$. However, since $0 < 1 < a$ as well, it must be that $0 \in B_5$, and hence $0 \in U$ since $B_5 \subset U$. As this clearly contradicts the definition of U, it has to be that U is not in fact in $\mathcal{T}_5$ so that $\mathcal{T}_3 \not\subset \mathcal{T}_5$.

Now consider the set $U = (-\infty, 0)$, which is clearly in $\mathcal{T}_5$ since it is a basis element. However, since $\mathbb{R} - U = [0, \infty)$ is clearly neither all of $\mathbb{R}$ nor finite, it follows that $U \notin \mathcal{T}_3$. This shows that $\mathcal{T}_5 \not\subset \mathcal{T}_3$, which completes the proof that the two are incomparable. □

Now, the fact that $\mathcal{T}_1 \subsetneq \mathcal{T}_2$ was shown in Lemma 13.4. All that remains to be shown is that $\mathcal{T}_2 \subsetneq \mathcal{T}_4$ since the rest of the relations follow from the transitivity of proper inclusion.

Proof. First consider any basis element $B_2 \in \mathcal{B}_2$ and any $x \in B_2$. Either B_2 is (a, b) or $(a, b) - K$ for $a < b$ so that $a < x < b$ with $x \notin K$. In the former case clearly the set $B_4 = (a, x]$ is in $\mathcal{B}_4$, $x \in B_4$, and $B_4 \subset B_2$. In the latter case we have the following:

Case: $x \leq 0$. Then here again $B_4 = (a, x]$ is in $\mathcal{B}_4$, $x \in B_4$, and $B_4 \subset B_2$ since $y \notin K$ for any $y \in B_4$ since then $a < y \leq x \leq 0$.

Case: $x > 0$. Then let n be the smallest positive integer where $n > 1/x$, which exists since $\mathbb{Z}_+$ has no upper bound and is well-ordered. It then follows that $0 < 1/n < x$ and there are no integers m such that $1/n < 1/m \leq x$. So let $a' = \max(a, 1/n)$ and set $B_4 = (a', x]$ so that, for any $y \in B_4$, both $a \leq a' < y \leq x < b$ and $1/n \leq a' < y < x$, and hence $y \in (a, b)$ and $y \notin K$. Therefore $y \in (a, b) - K = B_2$. Since y was arbitrary, this shows that $B_4 \subset B_2$, noting that also clearly $x \in B_4$ and $B_4 \in \mathcal{B}_4$.

Hence in any case it follows that $\mathcal{T}_2 \subset \mathcal{T}_4$ from Lemma 3.13.

Now let $x = -1$ and $B_4 = (-2, -1]$ so that clearly $x \in B_4$ and $B_4 \in \mathcal{B}_4$. Then let B_2 be any basis element in $\mathcal{B}_2$ that contains x. Then we have that B_2 is either (a, b) or $(a, b) - K$ where $a < x < b$ and $x \notin K$.

Case: $0 < b$. Then $a < x = -1 < 0 \leq b$ so that 0 is in both (a, b) and $(a, b) - K$ since clearly $0 \notin K$, and thus $0 \in B_2$. However, clearly $0 \notin (-2, -1] = B_4$.

Case: $0 \geq b$. Then $a < x < (x + b)/2 < b \leq 0$ so that $(x + b)/2 \in B_2$ since $(x + b)/2$ is not in K. Clearly also though $(x + b)/2 \notin (-2, x] = B_4$ since $x < (x + b)/2$.

Thus in either case we have that $B_2 \not\subset B_4$. This shows the negation of Lemma 13.3 so that $\mathcal{T}_4 \not\subset \mathcal{T}_2$. Hence $\mathcal{T}_2 \subsetneq \mathcal{T}_4$ as desired. □

It is perhaps a rather surprising fact that, though it has been shown that the K and lower limit topology are incomparable (Exercise 13.6), the K topology and the upper limit topology *are* comparable as was just shown.

Exercise 13.8

(a) Apply Lemma 13.2 to show that the countable collection

$$\mathcal{B} = \{(a,b) \mid a < b,\ a \text{ and } b \text{ rational}\}$$

is a basis that generates the standard topology on $\mathbb{R}$.

(b) Show that the collection

$$\mathcal{C} = \{[a,b) \mid a < b,\ a \text{ and } b \text{ rational}\}$$

is a basis that generates a topology different from the lower limit topology on $\mathbb{R}$.

Solution:

(a)

Proof. Let $\mathcal{T}$ be the standard topology on $\mathbb{R}$. First, clearly $\mathcal{B}$ is a collection of open sets of $\mathcal{T}$ since each element is a basis element in the standard basis (i.e. an open interval). Now consider any $U \in \mathcal{T}$ and any $x \in U$. Then there is a standard basis element $B' = (a', b')$ such that $x \in B'$ and $B' \subset U$ since $\mathcal{T}$ is generated by the standard basis. Then $a' < x < b'$ so that, since the rationals are order-dense in the reals (shown in Exercise 4.9d), there are rational a and b such that $a' < a < x < b < b'$. Let $B = (a, b)$ so that clearly $x \in B$, $B \subset B' \subset U$, and $B \in \mathcal{B}$. This shows that $\mathcal{B}$ is a basis for $\mathcal{T}$ by Lemma 13.2 since U and $x \in U$ were arbitrary. $\square$

(b)

Proof. First we must show that $\mathcal{C}$ is a basis at all. Clearly, for any $x \in \mathbb{R}$ we have that there is an element in $\mathcal{C}$ containing x, for example $[x, x+1)$. Now suppose that $C_1 = [a_1, b_1)$ and $C_2 = [a_2, b_2)$ are two elements of $\mathcal{C}$ and that $x \in C_1 \cap C_2$. Then obviously $a_1 \leq x < b_1$ and $a_2 \leq x < b_2$. Let $a = \max(a_1, a_2)$ and $b = \min(b_1, b_2)$ and $C = [a, b)$ so that clearly $C \in \mathcal{C}$. Also clearly $a \leq x < b$ since both $a_1 \leq x < b_1$ and $a_2 \leq x < b_2$, a is a_1 or a_2, and b is b_1 or b_2. Therefore C contains x. Now consider any $y \in C$ so that $a_1 \leq a \leq y < b \leq b_1$ and $a_2 \leq a \leq y < b \leq b_2$ and hence $y \in C_1$ and $y \in C_2$. This shows that $C \subset C_1 \cap C_2$ since y was arbitrary. By definition this suffices to show that $\mathcal{C}$ is a basis for a topology.

So let $\mathcal{T}$ be the topology generated by $\mathcal{C}$ and $\mathcal{T}_l$ be the lower limit topology. Now consider $U = [x, x+1)$ where x is any irrational number, for example $x = \pi$. Let C be any basis element in $\mathcal{C}$ containing x so that $C = [a, b)$ where a and b are rational. It must be that $a \neq x$ since a is rational but x is not. Also, since C contains x it has to be that $a \leq x$. So it has to be that $a < x$, but then $a \in C$ but $a \notin [x, x+1) = U$. This shows that C is not a subset of U. Hence we have shown

$$\exists x \in U \forall C \in \mathcal{C}\,(x \in C \Rightarrow C \not\subset U)$$
$$\exists x \in U \forall C \in \mathcal{C}\,(x \notin C \vee C \not\subset U)$$
$$\neg \forall x \in U \exists C \in \mathcal{C}\,(x \in C \wedge C \subset U)\ .$$

This shows that $U \notin \mathcal{T}$ by the definition of a generated topology. However, clearly we have that $U \in \mathcal{T}_l$ since it is a lower limit basis element. This suffices to show that $\mathcal{T}$ and $\mathcal{T}_l$ are different topologies. $\square$

§16 The Subspace Topology

Exercise 16.1

Show that if Y is a subspace of X and A is a subspace of Y, then the topology A inherits as a subspace of Y is the same as the topology it inherits as a subspace of X.

Solution:

Proof. Let $\mathcal{T}$ be the topology on X and $\mathcal{T}_Y$ be the subspace topology that Y inherits from X. Also let $\mathcal{T}_A$ and $\mathcal{T}'_A$ be the topologies that A inherits as a subspace of Y and X, respectively. Therefore we must show that $\mathcal{T}_A = \mathcal{T}'_A$. Now, by definition of subspace topologies we have that,

$$\mathcal{T}_Y = \{Y \cap U \mid U \in \mathcal{T}\} \qquad \mathcal{T}_A = \{A \cap U \mid U \in \mathcal{T}_Y\} \qquad \mathcal{T}'_A = \{A \cap U \mid U \in \mathcal{T}\} \; .$$

Now suppose that $W \in \mathcal{T}_A$ so that $W = A \cap V$ for some $V \in \mathcal{T}_Y$. Then we have that $V = Y \cap U$ for some $U \in \mathcal{T}$, and hence

$$W = A \cap V = A \cap (Y \cap U) = (A \cap Y) \cap U = A \cap U$$

since we have that $A \cap Y = A$ since $A \subset Y$. Since $U \in \mathcal{T}$ this clearly shows that $W \in \mathcal{T}'_A$ so that $\mathcal{T}_A \subset \mathcal{T}'_A$ since W was arbitrary.

Then, for any $W \in \mathcal{T}'_A$, we have that $W = A \cap U$ for some $U \in \mathcal{T}$. Let $V = Y \cap U$ so that clearly $V \in \mathcal{T}_Y$. Then as before we have that $A = A \cap Y$ since $A \subset Y$ so that

$$W = A \cap U = (A \cap Y) \cap U = A \cap (Y \cap U) = A \cap V \,,$$

and thus $W \in \mathcal{T}_A$ since $V \in \mathcal{T}_Y$. Since W was arbitrary this shows that $\mathcal{T}'_A \subset \mathcal{T}_A$, which completes the proof that $\mathcal{T}_A = \mathcal{T}'_A$. □

Exercise 16.2

If $\mathcal{T}$ and $\mathcal{T}'$ are topologies on X and $\mathcal{T}'$ is strictly finer than $\mathcal{T}$, what can you say about the corresponding subspace topologies on the subset Y of X?

Solution:

Let $\mathcal{T}_Y$ and $\mathcal{T}'_Y$ be the subspace topologies on Y corresponding to $\mathcal{T}$ and $\mathcal{T}'$, respectively. We claim that $\mathcal{T}_Y$ is finer than $\mathcal{T}_Y$ but not necessarily strictly finer.

Proof. First, we have that

$$\mathcal{T}_Y = \{Y \cap U \mid U \in \mathcal{T}\} \qquad \mathcal{T}'_Y = \{Y \cap U \mid U \in \mathcal{T}'\}$$

by the definition of subspace topologies. So for any $V \in \mathcal{T}_Y$ we have that $V = Y \cap U$ where $U \in \mathcal{T}$. Then also $U \in \mathcal{T}'$ since $\mathcal{T}'$ is finer than $\mathcal{T}$. This shows that $V \in \mathcal{T}'_Y$ since $V = Y \cap U$ where $U \in \mathcal{T}'$. Hence $\mathcal{T}'_Y$ is finer than $\mathcal{T}_Y$ since V was arbitrary.

To show that it is not necessarily strictly finer, consider the sets $X = \{a, b, c\}$ and $Y = \{a, b\}$ so that clearly $Y \subset X$. Consider also the topologies

$$\mathcal{T} = \{\varnothing, X, \{a, b\}\} \qquad \mathcal{T}' = \{\varnothing, X, \{a, b\}, \{c\}\}$$

on X so that clearly $\mathcal{T}'$ is strictly finer than $\mathcal{T}$. This results in the subspace topologies

$$\mathcal{T}_Y = \{\varnothing, Y\} \qquad \mathcal{T}'_Y = \{\varnothing, Y\} \; ,$$

which are clearly the same so that $\mathcal{T}'_Y$ is not strictly finer than $\mathcal{T}'_Y$, noting that it is technically still finer. However, if we instead have the topologies

$$\mathcal{T} = \{\varnothing, X, \{a, b\}\} \qquad \mathcal{T}' = \{\varnothing, X, \{a, b\}, \{b\}\}$$

then

$$\mathcal{T}_Y = \{\varnothing, Y\} \qquad\qquad \mathcal{T}'_Y = \{\varnothing, Y, \{b\}\}$$

so that $\mathcal{T}'_Y$ *is* strictly finer than $\mathcal{T}_Y$. Thus we can say nothing about the strictness of relation of the subspace topologies. □

Exercise 16.3

Consider the set $Y = [-1, 1]$ as a subspace of $\mathbb{R}$. Which of the following sets are open in Y? Which are open in $\mathbb{R}$?

$$\begin{aligned}
A &= \left\{x \mid \tfrac{1}{2} < |x| < 1\right\}, \\
B &= \left\{x \mid \tfrac{1}{2} < |x| \le 1\right\}, \\
C &= \left\{x \mid \tfrac{1}{2} \le |x| < 1\right\}, \\
D &= \left\{x \mid \tfrac{1}{2} \le |x| \le 1\right\}, \\
E &= \{x \mid 0 < |x| < 1 \text{ and } 1/x \notin \mathbb{Z}_+\}.
\end{aligned}$$

Solution:

Lemma 16.3.1. *If $a, b \in \mathbb{R}$ such that $0 \le a < b$ then the following are true:*

$$\{x \in \mathbb{R} \mid a < |x| < b\} = (-b, -a) \cup (a, b) \qquad \{x \in \mathbb{R} \mid a \le |x| \le b\} = [-b, -a] \cup [a, b]$$
$$\{x \in \mathbb{R} \mid a \le |x| < b\} = (-b, -a] \cup [a, b) \qquad \{x \in \mathbb{R} \mid a < |x| \le b\} = [-b, -a) \cup (a, b].$$

Proof. We prove only the first of these as the rest follow from nearly identical arguments. Let $A = \{x \in \mathbb{R} \mid a < |x| < b\}$ and $B = (-b, -a) \cup (a, b)$ so that we must show that $A = B$.

So consider $x \in A$ so that $a < |x| < b$. If $x \ge 0$ then $|x| = x$ so that $a < x < b$ and hence $x \in (a, b)$. If $x < 0$ then $|x| = -x$ so that $a < -x < b$, and thus $-a > x > -b$ so that $x \in (-b, -a)$. Thus in either case $x \in B$ so that $A \subset B$.

Now let $x \in B$ so that either $x \in (-b, -a)$ or $x \in (a, b)$. In the former case we have that $x < -a \le 0$ since $a \ge 0$ so that $|x| = -x$ and therefore

$$x \in (-b, -a) \Rightarrow -b < x < -a \Rightarrow b > -x = |x| > a \Rightarrow x \in A.$$

In the latter case we have that $x > a \ge 0$ so that $|x| = x$ and therefore

$$x \in (a, b) \Rightarrow a < x = |x| < b \Rightarrow x \in A.$$

This shows that $B \subset A$ since x was arbitrary, and thus $A = B$ as desired. □

Lemma 16.3.2. *Suppose that X is a topological space and $Y \subset X$ with the subspace topology. Then, if a set $U \subset Y$ is open in X, then it is also open in Y.*

Proof. So suppose that $U \subset Y$ is open in X. Then we have that $Y \cap U = U$ is also open in Y by the definition of the subspace topology. □

Main Problem.

First we claim that A is open in both $\mathbb{R}$ and Y.

Proof. We have from Lemma 16.3.1 that $A = (-1, -\frac{1}{2}) \cup (\frac{1}{2}, 1)$ which is clearly the union of basis elements so that A is open in $\mathbb{R}$. We also have that $A \subset Y$ so that A is open in Y by Lemma 16.3.2 since it is open in $\mathbb{R}$. $\square$

Next we claim that B is open in Y but not in $\mathbb{R}$.

Proof. By Lemmma 16.3.1 we have that $B = [-1, -\frac{1}{2}) \cup (\frac{1}{2}, 1]$. First, consider the sets $(-2, -\frac{1}{2})$ and $(\frac{1}{2}, 2)$, which are clearly both basis elements and therefore open in $\mathbb{R}$. We then have that $(-2, -\frac{1}{2}) \cap Y = [-1, -\frac{1}{2})$ and $(\frac{1}{2}, 2) \cap Y = (\frac{1}{2}, 1]$ so that these sets are open in Y by the definition of the subspace topology. Clearly then their union $B = [-1, -\frac{1}{2}) \cup (\frac{1}{2}, 1]$ is then also open in Y.

It is also easy to see that B is not open in $\mathbb{R}$. For example, $-1 \in B$ but for any basis element $B' = (a, b)$ containing -1 we have that $a < -1 < b$ so that $a < (a-1)/2 < -1 < b$ and hence $(a-1)/2 \in B'$. Clearly though $(a-1)/2 \notin B$ so that B' cannot be a subset of B. Thus suffices to show that B is not open by the definition of the topology of $\mathbb{R}$ generated by its basis. $\square$

We claim that C is open neither in $\mathbb{R}$ nor Y.

Proof. By Lemmma 16.3.1 we have that $C = (-1, -\frac{1}{2}] \cup [\frac{1}{2}, 1)$. If $\mathcal{B}$ is the standard basis on $\mathbb{R}$, then, by Lemma 16.1, the set $\mathcal{B}_Y = \{B \cap Y \mid B \in \mathcal{B}\}$ is a basis for the subspace Y. So consider the point $x = \frac{1}{2}$ and any basis element $B_Y \in \mathcal{B}_Y$ containing x. Then we have that $B_Y = Y \cap B_X$ for some basis element $B_X = (a, b)$ in $\mathcal{B}$, and thus $a < x < b$ since $x \in B_X$. Let $a' = \max(a, -\frac{1}{2})$ and set $y = (a' + x)/2$ so that

$$a \le a' < (a' + x)/2 = y < x < b\,,$$

and hence $y \in (a, b) = B_X$. Also we have

$$-1 < -\tfrac{1}{2} \le a' < (a' + x)/2 = y < x = \tfrac{1}{2} < 1$$

so that $y \in [-1, 1] = Y$. Therefore $y \in Y \cap B_X = B_Y$. However, since $-\frac{1}{2} < y < \frac{1}{2}$, clearly $y \notin C$ so that B_Y cannot be a subset of C. Since the basis element $B_Y \in \mathcal{B}_Y$ was arbitrary, this suffices to show that C cannot be open in Y since $\mathcal{B}_Y$ is a basis. Since also clearly $C \subset Y$, it follows from the contrapositive of Lemma 16.3.2 that C is not open in $\mathbb{R}$ either. $\square$

Next we claim that D is also not open in $\mathbb{R}$ or Y.

Proof. This follows from basically the same argument as the previous proof, again using the point $x = \frac{1}{2}$ to show that any basis element of Y that contains x cannot be a subset of D. $\square$

Lastly, we claim that E is open in both $\mathbb{R}$ and Y.

Proof. First, it is trivial to show that

$$E = \{x \in \mathbb{R} \mid 0 < |x| < 1\} - K = [(-1, 0) \cup (0, 1)] - K\,,$$

where we have used Lemma 16.3.1. Now consider any $x \in E$ so that $x \in (-1, 0) \cup (0, 1)$ and $x \notin K$. If $x \in (-1, 0)$ then clearly the basis element $(-1, 0)$ contains x and is a subset of E since $(-1, 0) \cap K = \varnothing$.

On the other hand, if $x \in (0, 1)$ then $x \notin K$ so that $1/x \notin \mathbb{Z}_+$. From this it follows from Exercise 4.9b that there is exactly one positive integer n such that $n < 1/x < n + 1$. We then have that $1/(n+1) < x < 1/n$. So let $B = (1/(n+1), 1/n)$ so that clearly $x \in B$, $B \cap K = \varnothing$, and B is a basis element of the standard topology on $\mathbb{R}$. Since $B \cap K = \varnothing$ and clearly $0 < 1/(n+1) < 1/n \le 1$, it also follows that $B \subset E$.

Hence in either case there is a basis element of $\mathbb{R}$ that contains x and is a subset of E. This suffices to show that E is open in $\mathbb{R}$. Since clearly $E \subset Y$, we also clearly have that E is open in Y by Lemma 16.3.2. $\square$

Exercise 16.4

A map $f : X \to Y$ is said to be an ***open map*** if for every open set U of X, the set $f(U)$ is open in Y. Show that $\pi_1 : X \times Y \to X$ and $\pi_2 : X \times Y \to Y$ are open maps.

Solution:

Proof. Suppose that U is an open subset of $X \times Y$. Consider any $x \in \pi_1(U)$ so that there is a $y \in Y$ such that $(x, y) \in U$. Then there is a basis element $A \times B$ of the product topology on $X \times Y$ where $(x, y) \in A \times B \subset U$. Then A and B are open sets of X and Y, respectively, since $A \times B$ is a basis element of the product topology. Clearly we have that $x \in A$ since $(x, y) \in A \times B$. Now, for any $x' \in A$, we have that $(x', y) \in A \times B$ so that $(x', y) \in U$. Hence $x' = \pi_1(x', y) \in \pi_1(U)$, which shows that $A \subset \pi_1(U)$ since x' was arbitrary. Then, since A is an open subset of X, there is a basis element A' where $x \in A' \subset A \subset \pi_1(U)$. This suffices to show that $\pi_1(U)$ is an open subset of X since x was arbitrary. An analogous argument shows that π_2 is also an open map. $\square$

Exercise 16.5

Let X and X' denote a single set in the topologies $\mathcal{T}$ and $\mathcal{T}'$, respectively; let Y and Y' denote a single set in the topologies $\mathcal{U}$ and $\mathcal{U}'$, respectively. Assume these sets are nonempty.

(a) Show that if $\mathcal{T}' \supset \mathcal{T}$ and $\mathcal{U}' \supset \mathcal{U}$, then the product topology on $X' \times Y'$ is finer than the product topology on $X \times Y$.

(b) Does the converse of (a) hold? Justify your answer.

Solution:

In what follows let $\mathcal{T}'_p$ and $\mathcal{T}_p$ denote the product topologies on $X' \times Y'$ and $X \times Y$, respectively.

(a)

Proof. Consider any $W \in \mathcal{T}_p$ and any $(x, y) \in W$, noting that obviously $W \subset X \times Y$. Then there is a basis element $U \times V$ of $\mathcal{T}_p$ such that $(x, y) \in U \times V$ and $U \times V \subset W$. By the definition of the product topology, we have that U and V are open sets in $\mathcal{T}$ and $\mathcal{U}$, respectively. Then we also have that $U \in \mathcal{T}'$ and $V \in \mathcal{U}'$ since $\mathcal{T} \subset \mathcal{T}'$ and $\mathcal{U} \subset \mathcal{U}'$. Hence $U \times V$ is also a basis element of $\mathcal{T}'_p$. Since we know that $(x, y) \in U \times V$, $U \times V \subset W$, and $(x, y) \in W$ was arbitrary, this suffices to show that W is an open subset of $X' \times Y'$ and hence $W \in \mathcal{T}'_p$. This in turn shows that $\mathcal{T}_p \subset \mathcal{T}'_p$ since W was arbitrary. $\square$

(b) We claim that the converse does not always hold.

Proof. As a counterexample consider $A = \{a, b, c, d\}$ so that clearly

$$\begin{aligned} \mathcal{T}' &= \{\varnothing, A, \{a, b\}, \{c, d\}\} \\ \mathcal{T} &= \{\varnothing, A, \{a, b\}, \{c, d\}, \{c\}, \{d\}, \{a, b, c\}, \{a, b, d\}\} \end{aligned}$$

are topologies on A. Clearly also $\mathcal{T}'$ is not finer than $\mathcal{T}$. Similarly let $B = \{1, 2, 3, 4\}$ so that

$$\begin{aligned} \mathcal{U}' &= \{\varnothing, B, \{1, 2\}, \{3, 4\}\} \\ \mathcal{U} &= \{\varnothing, B, \{1, 2\}, \{3, 4\}, \{3\}, \{4\}, \{1, 2, 3\}, \{1, 2, 4\}\} \end{aligned}$$

are topologies on B, also noting that clearly $\mathcal{U}'$ is not finer that $\mathcal{U}$. Now let $X = X' = \{a, b\}$ and $Y = Y' = \{1, 2\}$ so that clearly X and X' are in topologies $\mathcal{T}$ and $\mathcal{T}'$, respectively, and Y and Y' are in $\mathcal{U}$ and $\mathcal{U}'$, respectively.

Then the bases for the product topologies $\mathcal{T}_p$ on $X \times Y$ and $\mathcal{T}'_p$ on $X' \times Y'$ are then

$$\mathcal{B} = \{\varnothing, X \times Y\} \qquad \mathcal{B}' = \{\varnothing, X' \times Y'\} = \{\varnothing, X \times Y\} = \mathcal{B},$$

respectively, since there are no subsets of X in $\mathcal{T}$ or $\mathcal{T}'$ other than $\varnothing$ and X itself, and similarly no subsets of Y in $\mathcal{U}$ or $\mathcal{U}'$ other than $\varnothing$ and Y. Since their bases are the same, clearly $\mathcal{T}_p = \mathcal{T}'_p$ so that it is true that $\mathcal{T}'_p$ is finer than $\mathcal{T}_p$ (though not strictly so). □

Exercise 16.6

Show that the countable collection

$$\{(a, b) \times (c, d) \mid a < b \text{ and } c < d \text{ and } a, b, c, d \text{ are rational}\}$$

is a basis for $\mathbb{R}^2$.

Solution:

Proof. It was proven in Exercise 13.8a that the set

$$\mathcal{B} = \{(a, b) \mid a < b,\ a \text{ and } b \text{ rational}\}$$

is a basis for the standard topology on $\mathbb{R}$. It then follows that

$$\mathcal{D} = \{B \times C \mid B, C \in \mathcal{B}\}$$

is a basis for the standard topology on $\mathbb{R}^2$ by Theorem 15.1. Clearly we have

$$\mathcal{D} = \{(a, b) \times (c, d) \mid a < b \text{ and } c < d \text{ and } a, b, c, d \text{ are rational}\},$$

which shows the desired result. □

Exercise 16.7

Let X be an ordered set. If Y is a proper subset of X that is convex in X, does it follow that Y is an interval or a ray in X?

Solution:

We claim that Y is not always an interval or a ray in X.

Proof. As a counterexample consider $X = \mathbb{Q}$ and the proper subset $Y = \{x \in \mathbb{Q} \mid x^2 < 2\}$. We claim that Y is convex but not an interval or a ray.

First, consider $a, b \in Y$ where $a < b$, thus $a^2, b^2 < 2$. Also consider $x \in (a, b)$ so that $a < x < b$. If $x \geq 0$ then $0 \leq x < b$ so that $x^2 < b^2 < 2$. If $x < 0$ then $a < x < 0$ so that $2 > a^2 > x^2$. Thus in either case $x^2 < 2$ so that $x \in Y$. Since x was arbitrary, this shows that $(a, b) \subset Y$ so that Y is convex since a and b were arbitrary.

Now, clearly Y cannot be a ray with no lower bound since then there would be an x in the ray where $x < -2$ so that $x^2 > 4 > 2$ and hence $x \notin Y$. Similarly Y cannot be a ray with no upper bound since then the ray would contain an $x > 2$ so that $x^2 > 4 > 2$ and thus $x \notin Y$. So suppose that

$Y = [a, b]$ for some $a, b \in X = \mathbb{Q}$ where $a \leq b$. Now, it cannot be that $b^2 = 2$ since then $b = \sqrt{2}$, which is not rational. Similarly it cannot be that $a^2 = 2$ for the same reason.

Case: $b^2 < 2$. Then there is a rational p where $b < p < \sqrt{2}$ since the rationals are order-dense in the reals. Let $x = \max(0, p)$ so that $b < p \leq x$ and hence $x \notin [a, b]$. However, if $0 < p$ then $x = p$ so that $x^2 = p^2 < 2$, and if $0 \geq p$ then $x = 0$ so that $x^2 = 0 < 2$. Thus either way $x \in Y$ and $x \notin [a, b]$, which shows that Y cannot be $[a, b]$.

Case: $b^2 > 2$. Then $\sqrt{2} < b$ since $0 < 2 < b^2$. If $\sqrt{2} < a$ then clearly for any $x \in [a, b]$ we have that $0 < \sqrt{2} < a \leq x$ so that $2 < x^2$ and hence $x \notin Y$. If $a < \sqrt{2}$ then there is a rational p such that $a < \sqrt{2} < p < b$ since the rationals are order-dense in the reals. Hence $2 < p^2$ so that $p \notin Y$. Either way there is an $x \in [a, b]$ where $x \notin Y$ so that Y cannot be $[a, b]$.

Similar arguments show that neither $Y = (a, b)$, $Y = [a, b)$, nor $Y = (a, b]$ for $a, b \in X = \mathbb{Q}$ and $a < b$. Hence Y cannot be an interval. Thus Y is convex but neither an interval nor a ray in X. This shows the desired result. □

Exercise 16.8

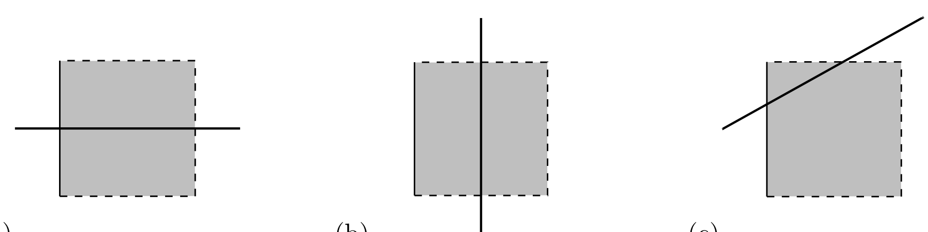
If L is a straight line in the plane, describe the topology L inherits as a subspace of $\mathbb{R}_l \times \mathbb{R}$ and as a subspace of $\mathbb{R}_l \times \mathbb{R}_l$. In each case it is a familiar topology.

Solution:

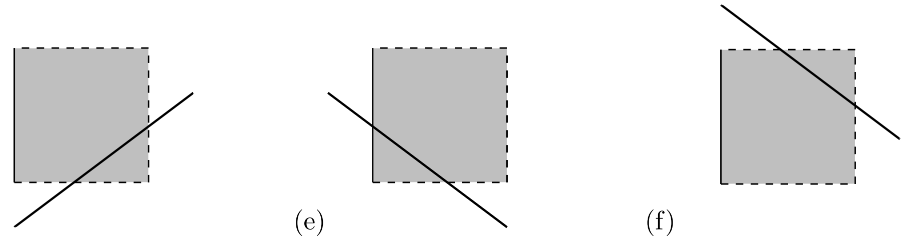
First, let $\mathbb{R}_u$denote the reals with the upper limit topology, with a basis containing all intervals $(a, b]$ for $a < b$. Also let $\mathbb{R}_d$denote the reals with the discrete topology, which can clearly be generated by a basis containing intervals $[a, b]$ for $a \leq b$. This is easy to see as $[a, a] = \{a\}$ is a basis element so that any subset of $\mathbb{R}$ can be considered a union of such basis elements. It is then easy to show that $\mathbb{R}_l$and $\mathbb{R}_u$are both strictly finer than the standard topology on $\mathbb{R}$ (this was shown in Lemma 13.4 for $\mathbb{R}_l$), but that $\mathbb{R}_l$and $\mathbb{R}_u$are incomparable. Clearly $\mathbb{R}_d$is strictly finer than both of these since it is the finest possible topology on $\mathbb{R}$.

Now, regarding the main problem, we do not yet have the tools show formally show how topologies on a line L compare to topologies on $\mathbb{R}$, so we will have to discuss this informally. We can see that, in some sense, a line L in the plane is like a copy of the real line so that we can discuss topologies on L is being in some sense the same as topologies on $\mathbb{R}$.

The product topology $\mathbb{R}_l \times \mathbb{R}$ is the topology is generated by the basis containing sets of the form $[a, b) \times (c, d)$ where $a < b$ and $c < d$ by Theorem 15.1. Then, for a line L in the plane, it can intersect such a basis element in a variety of ways, which are illustrated below:



(a) (b) (c)



(d) (e) (f)

Clearly the intersection of L and these basis elements results in some kind of interval on L. Such intervals then form the basis for the subspace topology on L by Lemma 16.1 since they are the intersection of L and a basis element in the superspace. Another point is that the orientation of the line L with regard to the way in which it is a copy of $\mathbb{R}$ is important. For example, in Figure (a) above, if L is oriented in the natural way with the negative reals on the left and the positive reals on the right, then the resulting intervals are of the form $[a, b)$, which would result in a topology like $\mathbb{R}_l$. The opposite orientation results in intervals of the form $(a, b]$ as basis elements, generating a topology like $\mathbb{R}_u$.

Now, for a line L such as that illustrated in Figure (a), every possible basis element of $\mathbb{R}_l \times \mathbb{R}$ that intersects L results the half-open intervals as described above depending on the orientation of L. This is not the case for all lines, however, and is dependent on its slope in the plane. For example, lines with positive slope can intersect basis elements as in Figure (c), which result in half open intervals $[a, b)$ (or $(a, b]$ depending on orientation), or they can intersect them as in Figure (d), which result in open intervals (a, b). However, since the topologies $\mathbb{R}_l$and $\mathbb{R}_u$are strictly finer than the standard topology, the subspace topology formed on L would be like these (which depends on orientation) rather than like the standard topology. Lastly, we note that, for any appropriate interval on the line L, we can clearly always find a basis element B in $\mathbb{R}_l \times \mathbb{R}$ such that the intersection of B with L is the interval. For this reason, these intervals form the basis elements of the topology on L.

With all these considerations in mind, we list the topologies on $\mathbb{R}$ that the subspace topologies on L are like based on line directions and orientations for product topologies $\mathbb{R}_l \times \mathbb{R}$ and $\mathbb{R}_l \times \mathbb{R}_l$:

L	$\mathbb{R}_l \times \mathbb{R}$	$\mathbb{R}_l \times \mathbb{R}_l$
$\rightarrow$	$\mathbb{R}_l$	$\mathbb{R}_l$
$\nearrow$	$\mathbb{R}_l$	$\mathbb{R}_l$
$\uparrow$	$\mathbb{R}$	$\mathbb{R}_l$
$\nwarrow$	$\mathbb{R}_u$	$\mathbb{R}_d$
$\leftarrow$	$\mathbb{R}_u$	$\mathbb{R}_u$
$\swarrow$	$\mathbb{R}_u$	$\mathbb{R}_u$
$\downarrow$	$\mathbb{R}$	$\mathbb{R}_u$
$\searrow$	$\mathbb{R}_l$	$\mathbb{R}_d$

We note that $\mathbb{R}$ simply denotes the standard topology.

Exercise 16.9

Show that the dictionary order topology on the set $\mathbb{R} \times \mathbb{R}$ is the same as the product topology on $\mathbb{R}_d \times \mathbb{R}$, where $\mathbb{R}_d$ denotes $\mathbb{R}$ in the discrete topology. Compare this topology with the standard topology on $\mathbb{R}^2$.

Solution:

In what follows let $\mathcal{T}_d$ denote the dictionary order topology on $\mathbb{R} \times \mathbb{R}$, and let $\mathcal{T}_p$ denote the product topology on $\mathbb{R}_d \times \mathbb{R}$. Also let $\prec$ denote the dictionary ordering of $\mathbb{R} \times \mathbb{R}$. First we show that $\mathcal{T}_d = \mathcal{T}_p$.

Proof. First we note that clearly the dictionary order on $\mathbb{R} \times \mathbb{R}$ has no largest or smallest elements so that, by definition, $\mathcal{T}_d$ has as basis elements intervals $((x, y), (x', y'))$, that is the set of all points $z \in \mathbb{R} \times \mathbb{R}$ where $(x, y) \prec z \prec (x', y')$. Clearly the set $\{\{x\} \mid x \in \mathbb{R}\}$ is a basis for $\mathbb{R}_d$. Hence, by Theorem 15.1, the set $\mathcal{B}_p = \{\{x\} \times (a, b) \mid x \in \mathbb{R} \text{ and } a < b\}$ is a basis for the product topology $\mathcal{T}_p$.

So consider any $(x, y) \in \mathbb{R} \times \mathbb{R}$ and any basis element $B_d = ((a, b), (a', b'))$ of $\mathcal{T}_d$ that contains (x, y). Hence $(a, b) \prec (x, y) \prec (a', b')$.

Case: $a = x$: Then since $(a, b) \prec (x, y)$, it has to be that $b < y$.

Case: $a = x = a'$. Then it also has to be that $y < b'$ since $(x, y) \prec (a', b')$. Then the set $B_p = \{x\} \times (b, b')$ is a basis element of $\mathcal{T}_p$ that contains (x, y) and is a subset of B_d.

Case: $a = x < a'$. Then it is easy to show that the set $B_p = \{x\} \times (b, y + 1)$ is a basis element of $\mathcal{T}_p$ that contains (x, y) and is a subset of B_d.

Case: $a < x$:

Case: $x = a'$. Then it has to be that $y < b'$ since $(x, y) \prec (a', b')$. Then it is easy to show that the set $B_p = \{x\} \times (y - 1, b')$ is a basis element of $\mathcal{T}_p$ that contains (x, y) and is a subset of B_d.

Case: $a = x < a'$. Then it is easy to show that the set $B_p = \{x\} \times (y - 1, y + 1)$ is a basis element of $\mathcal{T}_p$ that contains (x, y) and is a subset of B_d.

In every case and sub-case it follows from Lemma 13.3 that $\mathcal{T}_d \subset \mathcal{T}_p$.

Now suppose $(x, y) \in \mathbb{R} \times \mathbb{R}$ and $B_p = \{x\} \times (a, b)$ is a basis element of $\mathcal{T}_p$ containing (x, y). Also let B_d be the interval in the dictionary order $((x, a), (x, b))$, which is clearly a basis element of $\mathcal{T}_d$. It is then trivial to show that $B_p = B_d$ so that $x \in B_d \subset B_p$, which shows that $\mathcal{T}_p \subset \mathcal{T}_d$ by Lemma 13.3. This suffices to show that $\mathcal{T}_d = \mathcal{T}_p$ as desired. □

We now claim that this topology $\mathcal{T}_d = \mathcal{T}_p$ is strictly finer than the standard topology on $\mathbb{R} \times \mathbb{R}$. We denote the latter by simply $\mathcal{T}$.

Proof. Since it was just shown that $\mathcal{T}_d = \mathcal{T}_p$, it suffices to show that either one is strictly finer than the standard topology. It shall be most convenient to use the product topology $\mathcal{T}_p$. So first consider any $(x, y) \in \mathbb{R}^2$ and any basis element $B = (a, b) \times (c, d)$ of $\mathcal{T}$ containing (x, y). Hence $a < x < b$ and $c < y < d$. It is then trivial to show that the set $\{x\} \times (c, d)$, which is clearly a basis element of $\mathcal{T}_p$, contains (x, y) and is a subset of B. This shows that $\mathcal{T}_p$ is finer than $\mathcal{T}$ by Lemma 13.3.

To show that it is strictly finer, consider the point $(0, 0)$ and the set $B_p = \{0\} \times (-1, 1)$, which clearly contains $(0, 0)$ and is a basis element of $\mathcal{T}_p$. Now consider any basis element $B = (a, b) \times (c, d)$ of $\mathcal{T}$ that also contains $(0, 0)$. It then follows that $a < 0 < b$ and $c < 0 < d$. Consider then the point $x = (a + 0)/2 = a/2$ so that clearly $a < x < 0 < b$ and hence $x \in (a, b)$. Thus the point $(x, 0) \in B$, but also $(x, 0) \notin B_p$ since $x < 0$ so that $x \neq 0$. This shows that B cannot be a subset of B_p. Since B was an arbitrary basis element of $\mathcal{T}$, this shows that $\mathcal{T}$ is *not* finer than $\mathcal{T}_p$ by the negation of Lemma 13.3.

This suffices to show that $\mathcal{T}_p$ is strictly finer than $\mathcal{T}$ as desired. □

Exercise 16.10

Let $I = [0, 1]$. Compare the product topology on $I \times I$, the dictionary order topology on $I \times I$, and the topology $I \times I$ inherits as a subspace of $\mathbb{R} \times \mathbb{R}$ in the dictionary order topology.

Solution:

First, we assume that the product topology on $I \times I$ is the product of I with the order topology as this seems to be the standard when no topology is explicitly specified. Denote this product topology by $\mathcal{T}_p$. Let $\mathcal{T}_d$ denote the dictionary order topology on $I \times I$, and let $\mathcal{T}_s$ denote the subspace topology on $I \times I$ inherited as a subspace of $\mathbb{R} \times \mathbb{R}$ in the dictionary order topology. Lastly, let $\prec$ denote the dictionary order on $I \times I$ and $\mathbb{R} \times \mathbb{R}$. To avoid ambiguity we also use the notation $x \times y$ to denote the ordered pair (x, y) and reserve parentheses for open intervals.

First we claim that $\mathcal{T}_p$ and $\mathcal{T}_d$ are incomparable.

Proof. First, consider the point $0 \times 1 \in I \times I$ and $B_p = [0, 1/2) \times (1/2, 1]$, which is a basis element of $\mathcal{T}_p$ that clearly contains 0×1. Note that B_p is a basis element because $[0, 1/2)$ and $(1/2, 1]$ are both basis elements in the order topology on I since 0 and 1 are the smallest and largest elements of I, respectively. Now consider any interval $B_d = (a \times b, a' \times b')$ in the dictionary order on $I \times I$ that contains 0×1, which is of course a basis element of $\mathcal{T}_d$. Then we have that $a \times b$ and $a' \times b'$ are in $I \times I$ with $a \times b \prec 0 \times 1 \prec a' \times b'$. Hence $0 < a'$ or $0 = a'$ and $1 < b'$. As 1 is the largest element of I, the latter case is not possible so that it must be that $0 < a'$. Let $x = (0 + a')/2 = a'/2$ so that clearly $0 < x < a'$. Then we have that $a \times b' \prec x \times 0 \prec a' \times b'$ so that the point $x \times 0$ is in B_d. However, clearly $0 \notin (1/2, 1]$ so that $x \times 0 \notin B_p$. This shows that B_d cannot be a subset of B_p.

Here we note that, in the dictionary order on $I \times I$, the smallest element is 0×0 while the largest is 1×1. With this in mind, the above argument for an open interval also applies to the half-open intervals $[0 \times 0, a \times b)$ and $(a \times b, 1 \times 1]$, which are of course also basis elements of $\mathcal{T}_d$. This then shows that $\mathcal{T}_d$ is not finer than $\mathcal{T}_p$ by the negation of Lemma 13.3.

Now consider the point $0 \times 1/2$ and the interval $B_d = (0 \times 0, 0 \times 1)$ in the dictionary ordering, which is therefore a basis element of $\mathcal{T}_d$, and clearly also contains $0 \times 1/2$. Consider also any basis element $B_p = A \times B$ of $\mathcal{T}_p$ that contains $0 \times 1/2$. Since $0 \in A$ and A must be a basis element of the order topology on I, it has to be that $A = [0, a)$ for some $0 < a \leq 1$. Then let $x = (0 + a)/2 = a/2$ so that $0 < x < a$, and thus $x \in A$. Then, since $1/2 \in B$ (since $0 \times 1/2 \in A \times B$), we have that $x \times 1/2 \in A \times B = B_p$ as well. However, we also clearly have that $0 \times 1 \prec x \times 1/2$ since $0 < x$ so that $x \times 1/2 \notin (0 \times 0, 0 \times 1) = B_d$. This shows that B_p cannot be a subset of B_d. As B_p was an arbitrary basis element of $\mathcal{T}_p$, this shows by the negation of Lemma 13.3 that $\mathcal{T}_p$ is not finer than $\mathcal{T}_d$.

This suffices to show that $\mathcal{T}_d$ and $\mathcal{T}_p$ are incomparable. $\square$

Next we claim that $\mathcal{T}_s$ is strictly finer than $\mathcal{T}_p$.

Proof. Consider any $x \times y \in I \times I$ and suppose that $B_p = A \times B$ is a basis element of $\mathcal{T}_p$ that contains $x \times y$. First suppose that $A = (a, b)$ and $B = (c, d)$ so that of course $a, b, c, d \in I$, $a < x < b$, and $c < y < d$. It is then trivial to show that the interval $B_s = (x \times c, x \times d)$ in the dictionary order also contains $x \times y$, is a basis element of $\mathcal{T}_s$ (since $B_s \subset I \times I$ so that $B_s \cap (I \times I) = B_s$), and is a subset of B_p. A similar argument can be made if A is an interval of the form $[0, a)$ or $(a, 1]$. If $B = (c, 1]$ and A is still (a, b), then let X be the interval $(x \times c, x \times 2)$ in the dictionary order so that we have $B_s = X \cap (I \times I) = \{x\} \times (c, 1]$ is a basis element of $\mathcal{T}_s$ that contains $x \times y$ and is a subset of B_p. A similar argument applies if $B = [0, d)$ and/or when the interval A is half-open. This shows that $\mathcal{T}_s$ is finer than $\mathcal{T}_p$ by Lemma 13.3.

The argument above that shows that $\mathcal{T}_p$ is not finer than $\mathcal{T}_d$ using the negation of Lemma 13.3 applies equally well to show that $\mathcal{T}_p$ is not finer than $\mathcal{T}_s$. This of course suffices to show the desired result that $\mathcal{T}_s$ is strictly finer than $\mathcal{T}_p$. $\square$

Lastly, we claim that $\mathcal{T}_s$ is also strictly finer than $\mathcal{T}_d$.

Proof. First consider any point $x \times y$ in $I \times I$ and let B_d be a basis element of $\mathcal{T}_d$ that contains $x \times y$ so that it is some kind of interval with endpoints $a \times b$ and $a' \times b'$ in $I \times I$. We note here that, since $\mathbb{R} \times \mathbb{R}$ has no smallest or largest elements, basis elements of the dictionary order topology there can only be open intervals. Now, if B_d is an open interval in $I \times I$ then clearly then the same interval $B = (a \times b, a' \times b')$ is a basis element in the dictionary order topology of $\mathbb{R} \times \mathbb{R}$, though though the two intervals can in general be different sets. For example the interval $(0 \times 0, 1 \times 1)$ in $\mathbb{R} \times \mathbb{R}$ contains the point 0×100 whereas the same interval in $I \times I$ does not since $0 \times 100 \notin I \times I$. It is, however, trivial to show that $B \cap (I \times I) = B_d$ so that B_d is basis element of $\mathcal{T}_s$.

If we have that B_d is the half-open interval $[0 \times 0, a' \times b')$ then let $B = (0 \times -1, a' \times b')$, which is clearly a basis element of the dictionary order topology on $\mathbb{R} \times \mathbb{R}$. It is then easy to see that $B \cap (I \times I) = B_d$ again so that it is a basis element of $\mathcal{T}_s$. If B_d is the half-open interval $(a \times b, 1 \times 1]$, then the open interval $(a \times b, 1 \times 2)$ is a basis element of the dictionary order topology on $\mathbb{R} \times \mathbb{R}$ and has the same result. Hence in all cases B_d is also a basis element of $\mathcal{T}_s$, and that it trivially is a subset of itself, and it contains $x \times y$. This shows that $\mathcal{T}_s$ is finer than $\mathcal{T}_d$ by Lemma 13.3.

To show that it is strictly finer, consider the point 0×1 and the open interval $B = (0 \times 0, 0 \times 2)$, which is clearly a basis element of the dictionary order topology in $\mathbb{R} \times \mathbb{R}$. It is then easy to prove that $B_s = B \cap (I \times I) = \{0\} \times (0, 1] = (0 \times 0, 0 \times 1]$ so that B_s is a basis element of $\mathcal{T}_s$. Now consider any basis element B_d of $\mathcal{T}_d$ that contains 0×1 so that B_d is some type of dictionary-order interval with endpoints $a \times b$ and $a' \times b'$, both in $I \times I$. The only way the interval can be closed above is if $a' \times b' = 1 \times 1$, in which case clearly $1 \times 1 \in B_d$ but $1 \times 1 \notin B_s$. So assume that it is open above so that $0 \times 1 \prec a' \times b'$, and hence either $0 < a'$ or $0 = a'$ and $1 < b'$. The latter case cannot be since 1 is the largest element of I and $b' \in I$. Therefore it has to be that $0 < a'$. So let $x = (0 + a')/2 = a'/2$ so that $0 < x < a'$ and thus $0 \times 1 \prec x \times 0 \prec a' \times b'$. From this it follows that $x \times 0$ is in B_d. However, clearly $x \times 0 \notin B_s$ since $0 \times 1 \prec x \times 0$.

Hence in any case we have shown that, while they both contain 0×1, B_d cannot be a subset of B_s. Since B_d was an arbitrary basis element, this shows that $\mathcal{T}_d$ is not finer than $\mathcal{T}_s$ by the negation of Lemma 13.3. This shows the desired result that $\mathcal{T}_s$ is strictly finer than $\mathcal{T}_d$. □

§17 Closed Sets and Limit Points

Exercise 17.1

Let $\mathcal{C}$ be a collection of subsets of the set X. Suppose that $\varnothing$ and X are in C, and that finite unions and arbitrary intersections of elements of $\mathcal{C}$ are in $\mathcal{C}$. Show that the collection

$$\mathcal{T} = \{X - C \mid C \in \mathcal{C}\}$$

is a topology on X.

Solution:

Proof. First, clearly $\varnothing$ and X are in $\mathcal{T}$ since $\varnothing = X - X$ and $X = X - \varnothing$ and both X and $\varnothing$ are in $\mathcal{C}$. This shows the first defining property of a topology.

Now consider an arbitrary sub-collection $\mathcal{A}$ of $\mathcal{T}$. Then, for each $A \in \mathcal{A}$, we have that $A = X - B$ for some $B \in \mathcal{C}$ since also $A \in \mathcal{T}$. So let $\mathcal{B} = \{B \in \mathcal{C} \mid X - B \in \mathcal{A}\}$. Then we have that

$$\bigcup \mathcal{A} = \bigcup_{A \in \mathcal{A}} A = \bigcup_{B \in \mathcal{B}} (X - B) = X - \bigcap_{B \in \mathcal{B}} B = X - \bigcap \mathcal{B}$$

by DeMorgan's law. By the definition of $\mathcal{C}$ we have that $\bigcap \mathcal{B} \in \mathcal{C}$ since it is an arbitrary intersection of elements of $\mathcal{C}$. It then follows that $\bigcup \mathcal{A} = X - \bigcap \mathcal{B}$ is in $\mathcal{T}$ by definition. This shows the second defining property of a topology.

Lastly, suppose that $\mathcal{A}$ is a nonempty finite sub-collection of $\mathcal{T}$, which of course can be expressed as $\mathcal{A} = \{A_k \mid k \in \{1, \ldots, n\}\}$ for some positive integer n. Then, again we have that that $A_k = X - B_k$ for some $B_k \in \mathcal{C}$ for all $k \in \{1, \ldots, n\}$ since $A_k \in \mathcal{T}$. Then we have

$$\bigcap \mathcal{A} = \bigcap_{k=1}^{n} A_k = \bigcap_{k=1}^{n} (X - B_k) = X - \bigcup_{k=1}^{n} B_k$$

by DeMorgan's law. Then clearly $\bigcup_{k=1}^{n} B_k$ is in $\mathcal{C}$ by definition since it is a finite union of elements of $\mathcal{C}$. It then follows that $\bigcap \mathcal{A} = X - \bigcup_{k=1}^{n} B_k$ is in $\mathcal{T}$ by definition. Since $\mathcal{A}$ was an arbitrary finite sub-collection, this shows the third defining property of a topology. Hence $\mathcal{T}$ is a topology by definition. □

Exercise 17.2

Show that if A is closed in Y and Y is closed in X, then A is closed in X.

Solution:

Proof. Since A is closed in Y, it follows from Theorem 17.2 that $A = B \cap Y$ where B is some closed set in X. Hence by definition $X - B$ is open in X. Also, since Y is closed in X, we have that $X - Y$ is open in X by definition. We then have

$$X - A = X - (B \cap Y) = (X - B) \cup (X - Y)$$

by DeMorgan's law. Since both $X - B$ and $X - Y$ are open in X, clearly their union must also be open since we are in a topological space. Hence $X - A$ is open in X so that A is closed in X by definition. □

Exercise 17.3

Show that if A is closed in X and B is closed in Y, then $A \times B$ is closed in $X \times Y$.

Solution:

Lemma 17.3.1. *If X, Y, A, and B are sets then $X \times Y - A \times B = (X - A) \times Y \cup X \times (Y - B)$.*

Proof. We show this via logical equivalences:

$$\begin{aligned}
(x, y) \in X \times Y - A \times B &\Leftrightarrow (x, y) \in X \times Y \wedge (x, y) \notin A \times B \\
&\Leftrightarrow (x \in X \wedge y \in Y) \wedge \neg(x \in A \wedge y \in B) \\
&\Leftrightarrow (x \in X \wedge y \in Y) \wedge (x \notin A \vee y \notin B) \\
&\Leftrightarrow (x \in X \wedge y \in Y \wedge x \notin A) \vee (x \in X \wedge y \in Y \wedge y \notin B) \\
&\Leftrightarrow (x \in X - A \wedge y \in Y) \vee (x \in X \wedge y \in Y - B) \\
&\Leftrightarrow (x, y) \in (X - A) \times Y \vee (x, y) \in X \times (Y - B) \\
&\Leftrightarrow (x, y) \in (X - A) \times Y \cup X \times (Y - B)
\end{aligned}$$

as desired. □

Main Problem.

Proof. Since A is closed we have that $X - A$ is open in X. Since also Y itself is open in Y, we have that $(X - A) \times Y$ is a basis element in the product topology by definition, and is therefore obviously open. An analogous argument shows that $X \times (Y - B)$ is also open in the product topology since B is closed in Y. Hence their union is also open in the product topology, but by Lemma 17.3.1 we have

$$(X - A) \times Y \cup X \times (Y - B) = X \times Y - A \times B$$

so that $X \times Y - A \times B$ is also open in the product topology. It then follows by definition that $A \times B$ is closed as desired. $\square$

Exercise 17.4

Show that if U is open in X and A is closed in X, then $U - A$ is open in X, and $A - U$ is closed in X.

Solution:

Lemma 17.4.1. *If A, B, and C are sets then $A - (B - C) = (A - B) \cup (A \cap C)$.*

Proof. We show this by a sequence of logical equivalences:

$$\begin{aligned} x \in A - (B - C) &\Leftrightarrow x \in A \wedge x \notin B - C \\ &\Leftrightarrow x \in A \wedge \neg(x \in B \wedge x \notin C) \\ &\Leftrightarrow x \in A \wedge (x \notin B \vee x \in C) \\ &\Leftrightarrow (x \in A \wedge x \notin B) \vee (x \in A \wedge x \in C) \\ &\Leftrightarrow x \in A - B \vee x \in A \cap C \\ &\Leftrightarrow x \in (A - B) \cup (A \cap C) \end{aligned}$$

as desired. $\square$

Corollary 17.4.2. *If $A \subset X$ and $B = X - A$, then $A = X - B$.*

Proof. By Lemma 17.4.1, we have that

$$X - B = X - (X - A) = (X - X) \cup (X \cap A) = \varnothing \cup (X \cap A) = X \cap A = A$$

since $A \subset X$. $\square$

Main Problem.

Proof. First, since A is closed in X, we have that $B = X - A$ is open in X, and it follows from Corollary 17.4.2 that $A = X - B$. Then we have that

$$U - A = U - (X - B) = (U - X) \cup (U \cap B)$$

by Lemma 17.4.1. Since $U \subset X$, it follows that $U - X = \varnothing$, and hence

$$U - A = \varnothing \cup (U \cap B) = U \cap B\,.$$

Then, since both U and B are open, their intersection is as well and therefore $U - A$ is open.

Next, we have by Lemma 17.4.1

$$X - (A - U) = (X - A) \cup (X \cap U) = B \cup (X \cap U) = B \cup U.$$

since $U \subset X$ so that $X \cap U = U$. Since both B and U are open, clearly their union is as well and hence $X - (A - U)$ is open. This of course means that $A - U$ is closed by definition. $\square$

Exercise 17.5

Let X be an ordered set in the order topology. Show that $\overline{(a,b)} \subset [a,b]$. Under what conditions does equality hold?

Solution:

Proof. First, the closed interval $[a,b]$ is closed (hence why it is called such!) because clearly its compliment is

$$X - [a,b] = (-\infty, a) \cup (b, \infty)$$

and we know that open rays are always open so that their union is as well. Clearly also $[a,b]$ contains (a,b). Hence $[a,b]$ is a closed set containing (a,b). Since $\overline{(a,b)}$ is defined as the intersection of closed sets that contain (a,b) clearly we have that $\overline{(a,b)} \subset [a,b]$ as desired. □

The conditions required for equality are such that $[a,b]$ is also a subset of $\overline{(a,b)}$ and, in particular both a and b must be in $\overline{(a,b)}$. Since clearly $a,b \notin (a,b)$, it has to be that they are both limit points of (a,b). This is equivalent to the condition that a has no immediate successor and b no immediate predecessor. We show only the first of these since the second is analogous.

Proof. ($\Rightarrow$) We show the contrapositive of this. So suppose that a *does* have an immediate successor c. Then the open ray $(-\infty, c)$ is an open set that contains a but does not intersect (a,b). This is easy to see, because if they did intersect, there would be an $x \in (a,b)$ where also $x \in (-\infty, c)$. From these it follows that $a < x < c$, which contradicts the fact that c is the immediate successor of a. Hence by definition a is not a limit point of (a,b).

($\Leftarrow$) Suppose that a is not a limit point of (a,b). Then there is an open set U containing a that does not intersect (a,b). From this it follows that there is a basis element B containing a such that $B \subset U$, and thus B also cannot intersect (a,b) (as, if it did, then so would U). Suppose that B is the open interval (c,d) so that $c < a < d$. It also must be that $d < b$ for otherwise, for any element of x of (a,b), we would have $c < a < x < b \leq d$ so that $x \in (c,d) = B$ and B and (a,b) would not be disjoint. We claim that d is the immediate successor of a. If this is not the case then there would be an x such that $c < a < x < d$ and hence $x \in (c,d) = B$. Also $a < x < d < b$ so that also $x \in (a,b)$. Therefore B and (a,b) would not be disjoint. Similar arguments can be made if B are other types of basis element in the order topology. (Actually B cannot be of the form $(e,f]$ for largest element f of X since then any element of (a,b) would also be in B and they would not be disjoint.) □

It is also worth noting that the Hausdorff axiom (and therefore also the T_1 axiom since it is implied by the Hausdorff axiom) is not sufficient for general equivalence of $[a,b]$ and $\overline{(a,b)}$. For example the order topology on $\mathbb{Z}$ results in the discrete topology so that every subset is both open and closed. Thus for any pair x_1, x_2 in $\mathbb{Z}$, the sets $\{x_1\}$ and $\{x_2\}$ are neighborhoods of x_1 and x_2, respectively, that are disjoint. This shows that this topology is a Hausdorff space. However, the fact that a has an immediate successor in $\mathbb{Z}_+$ is sufficient to show that $[a,b] \neq \overline{(a,b)}$ per what was just shown above.

Exercise 17.6

Let A, B, and A_α denote subsets of a space X. Prove the following:

(a) If $A \subset B$, then $\overline{A} \subset \overline{B}$.

(b) $\overline{A \cup B} = \overline{A} \cup \overline{B}$.

(c) $\overline{\bigcup A_a} \supset \bigcup \overline{A}_\alpha$; give an example where equality fails.

Solution:

(a)

Proof. Suppose that $A \subset B$ and consider any $x \in \overline{A}$. Consider any neighborhood U of x so that U intersects A by Theorem 17.5a. Hence there is a point $y \in U \cap A$ so that $y \in U$ and $y \in A$. But then clearly $y \in B$ also since $A \subset B$. Therefore $y \in U \cap B$ so that U intersects B. Since U was an arbitrary neighborhood of x, this shows that $x \in \overline{B}$, again by Theorem 17.5a. This of course shows that $\overline{A} \subset \overline{B}$ as desired since x was arbitrary. $\square$

(b)

Proof. ($\subset$) We show this by contrapositive. So suppose that $x \notin \overline{A} \cup \overline{B}$. Then clearly $x \notin \overline{A}$ and $x \notin \overline{B}$. Thus, by Theorem 17.5a, there is an open set U_A such that U_A does not intersect A, and likewise an open U_B that does not intersect B. Let $U = U_A \cap U_B$, which is clearly open since U_A and U_B are. We also note that U contains x since both U_A and U_B do. Then it must be that U does not intersect A since, if it did, then U_A would also intersect A since $U \subset U_A$. Similarly, U cannot intersect B. Thus, for all $y \in U$, $y \notin A$ and $y \notin B$. This is logically equivalent to saying that there is no $y \in U$ where $y \in A$ or $y \in B$, therefore there is no $y \in U$ where $y \in A \cup B$. Hence U and $A \cup B$ do not intersect. Since U is open and contains x, this shows that $x \notin \overline{A \cup B}$, again by Theorem 17.5a. Therefore, by contrapositive, $x \in \overline{A \cup B}$ implies that $x \in \overline{A} \cup \overline{B}$ so that $\overline{A \cup B} \subset \overline{A} \cup \overline{B}$.

($\supset$) Consider any $x \in \overline{A} \cup \overline{B}$ and any neighborhood U of x. If $x \in \overline{A}$ then U intersects A by Theorem 17.5a. Hence there is a $y \in U \cap A$ so that $y \in U$ and $y \in A$. Then clearly $y \in A \cup B$ so that y is also in $U \cap (A \cup B)$. Hence U intersects $A \cup B$. An analogous argument shows that this is also true if $x \in \overline{B}$ instead. Since U was an arbitrary neighborhood, this shows that $x \in \overline{A \cup B}$ by Theorem 17.5a. Hence $\overline{A} \cup \overline{B} \subset \overline{A \cup B}$ since x was arbitrary. $\square$

(c)

Proof. Consider any $x \in \bigcup \overline{A}_\alpha$ so that there is a particular β where $x \in \overline{A}_\beta$. Suppose that U is any open set containing x so that U intersects A_β by Theorem 17.5a since $x \in \overline{A}_\beta$. Then clearly U also intersects $\bigcup A_\alpha$ since $A_\beta \subset \bigcup A_\alpha$. Since U was an arbitrary open set containing x, this shows that $x \in \overline{\bigcup A_\alpha}$ by Theorem 17.5a. This shows that $\bigcup \overline{A}_\alpha \subset \overline{\bigcup A_\alpha}$ since x was arbitrary, which is of course the desired result. $\square$

As an example where equality fails, consider the standard topology on $\mathbb{R}$ and the sets $A_n = (1/n, 2]$ for $n \in \mathbb{Z}_+$. It is then trivial to show that $\bigcup A_n = (0, 2]$ so that clearly 0 is a limit point of $\bigcup A_n$, and hence $0 \in \overline{\bigcup A_n}$. However, for any $n \in \mathbb{Z}_+$, the open interval $(-1, 1/n)$ is clearly an open set containing 0 that is disjoint from $(1/n, 2] = A_n$. This shows that $0 \notin \overline{A}_n$ for every $n \in \mathbb{Z}_+$ by Theorem 17.5a, from which it follows that $0 \notin \bigcup \overline{A}_n$. Hence $\overline{\bigcup A_n}$ is not a subset of $\bigcup \overline{A}_n$ and thus $\overline{\bigcup A_n} \neq \bigcup \overline{A}_n$.

Exercise 17.7

Criticize the following "proof" that $\overline{\bigcup A_\alpha} \subset \bigcup \overline{A}_\alpha$: if $\{A_\alpha\}$ is a collection of sets in X and if $x \in \overline{\bigcup A_\alpha}$, then every neighborhood U of x intersects $\bigcup A_\alpha$. Thus U must intersect some A_α, so that x must belong to the closure of some A_α. Therefore, $x \in \bigcup \overline{A}_\alpha$.

Solution:

The problem with this "proof" is that, just because every neighborhood U intersects *some* A_α, it does not mean that *every* U intersects a single A_α, which is what is required for x to be in $\overline{A}_\alpha$. This is illustrated in the counterexample above at the end of Exercise 17.6c. There, every neighborhood of 0 clearly intersects *some* set $A_n = (1/n, 2]$, but, for any given $n \in \mathbb{Z}_+$, not every neighborhood of 0 intersects A_n, for example the neighborhood $(-1, 1/n)$ does not.

Exercise 17.8

Let A, B, and A_α denote subsets of a space X. Determine whether the following equations hold; if an equality fails, determine whether one of the inclusions $\supset$ or $\subset$ holds.

(a) $\overline{A \cap B} = \overline{A} \cap \overline{B}$.

(b) $\overline{\bigcap A_\alpha} = \bigcap \overline{A}_\alpha$.

(c) $\overline{A - B} = \overline{A} - \overline{B}$.

Solution:

(a) We claim that $\overline{A \cap B} \subset \overline{A} \cap \overline{B}$ but equality is not always true.

Proof. Consider any $x \in \overline{A \cap B}$ and any open set U containing x. Then by, Theorem 17.5a, U intersects $A \cap B$, from which it immediately follows that U intersects both A and B. However, since U was an arbitrary neighborhood of x, it follows from Theorem 17.5a again that x is in both $\overline{A}$ and $\overline{B}$. Hence $x \in \overline{A} \cap \overline{B}$, which shows that $\overline{A \cap B} \subset \overline{A} \cap \overline{B}$ since x was arbitrary.

Now consider the standard topology on $\mathbb{R}$ with $A = [-1, 0)$ and $B = (0, 1]$. As these are clearly disjoint, we have that $A \cap B = \varnothing$ so that $\overline{A \cap B} = \varnothing$ also. However, since we also clearly have that $\overline{A} = [-1, 0]$ and $\overline{B} = [0, 1]$, it follows that $\overline{A} \cap \overline{B} = \{0\}$. Thus clearly $\overline{A \cap B} = \varnothing \neq \{0\} = \overline{A} - \overline{B}$ as desired. □

(b) We again claim that $\overline{\bigcap A_\alpha} \subset \bigcap \overline{A}_\alpha$ but that equality is not generally true.

Proof. Consider any $x \in \overline{\bigcap A_\alpha}$ and any open set U of x. Then, by Theorem 17.5a, U intersects $\bigcap A_\alpha$ so that, for any particular A_β, U intersects A_β. This shows that $x \in \overline{A}_\beta$ by Theorem 17.5a so that $x \in \overline{A}_\alpha$ for every α since β was arbitrary. Hence $x \in \bigcap \overline{A}_\alpha$, which shows that $\overline{\bigcap A_\alpha} \subset \bigcap \overline{A}_\alpha$ since x was arbitrary.

As in part (a), equality fails if we have $A_1 = [-1, 0)$ and $A_2 = (0, 1]$ in the standard topology on $\mathbb{R}$. By the same argument as in part (a) it follows that $\overline{\bigcap_{n=1}^{2} A_n} = \varnothing \neq \{0\} = \bigcap_{n=1}^{2} \overline{A}_n$. □

(c) Here we claim that $\overline{A - B} \supset \overline{A} - \overline{B}$ but that the converse does not always hold.

Proof. Consider any $x \in \overline{A} - \overline{B}$ and any open set U containing x. Then $x \in \overline{A}$ so that every open set containing x intersects A by Theorem 17.5a. Also $x \notin \overline{B}$ so that there is an open set V containing x that does not intersect B, also by Theorem 17.5a. Let $W = U \cap V$ so that W contains x since both $x \in U$ and $x \in V$. Now, since W is also an open set containing x, W intersects A so that there is a $y \in W$ where also $y \in A$. It also cannot be that $y \in B$ since we have $y \in W \subset V$ so that then V would intersect B. Therefore $y \in A - B$. Also we have $y \in W \subset U$ so that also $y \in U$. Hence U intersects $A - B$, which shows that $x \in \overline{A - B}$ by Theorem 17.5a since U was an arbitrary neighborhood of x. Therefore $\overline{A - B} \supset \overline{A} - \overline{B}$ as desired since x was arbitrary.

As a counterexample to equality, consider the standard topology on $\mathbb{R}$ with $A = [0, 2]$ and $B = (1, 3]$. Then clearly $\overline{A} = A = [0, 2]$ and $\overline{B} = [1, 3]$, from which it is easily shown that $\overline{A} - \overline{B} = [0, 1)$. But we also have $A - B = [0, 1]$ so that obviously $\overline{A - B} = [0, 1]$ as well. Therefore $\overline{A - B} = [0, 1] \neq [0, 1) = \overline{A} - \overline{B}$ as desired. □

Exercise 17.9

Let $A \subset X$ and $B \subset Y$. Show that in the space $X \times Y$,

$$\overline{A \times B} = \overline{A} \times \overline{B}\,.$$

Solution:

Proof. ($\subset$) Consider $(x, y) \in \overline{A \times B}$. Also suppose that U and V are any open sets in X and Y, respectively, that contain x and y, respectively. Then $U \times V$ is a basis element of the product topology on $X \times Y$, by definition, that contains (x, y). It then follows from Theorem 17.5b that $U \times V$ intersects $A \times B$ and hence there is a point $(w, z) \in U \times V$ where also $(w, z) \in A \times B$. Then $w \in U$ and $w \in A$ so that U intersects A, and hence $x \in \overline{A}$ by Theorem 17.5a since U was an arbitrary neighborhood of x. An analogous argument shows that $y \in \overline{B}$. Therefore $(x, y) \in \overline{A} \times \overline{B}$ so that $\overline{A \times B} \subset \overline{A} \times \overline{B}$ since x was arbitrary.

($\supset$) Now suppose that (x, y) is any point in $\overline{A} \times \overline{B}$ so that $x \in \overline{A}$ and $y \in \overline{B}$. Suppose also that $U \times V$ is any basis element of $X \times Y$ that contains (x, y) so that by definition U and V are open in X and Y, respectively. Since $x \in \overline{A}$ and U is an open set where $x \in U$, it follows from Theorem 17.5a that U intersects A. Thus there is $w \in U$ where $w \in A$ as well. An analogous argument shows that V intersects B so that there is a $z \in V$ where also $z \in B$. We therefore have that $(w, z) \in U \times V$ and $(w, z) \in A \times B$ so that $U \times V$ intersects $A \times B$. Since $U \times V$ was an arbitrary basis element containing (x, y), it follows from Theorem 17.5b that $(x, y) \in \overline{A \times B}$. This shows that $\overline{A} \times \overline{B} \subset \overline{A \times B}$ since the point (x, y) was arbitrary. $\square$

Exercise 17.10

Show that every order topology is Hausdorff.

Solution:

Proof. Suppose that X is an ordered set with the order topology. Consider a pair of distinct points x_1 and x_2 in X. Since X is an order, x_1 and x_2 must be comparable since they are distinct, so we can assume that $x_1 < x_2$ without loss of generality.

Case: x_2 is the immediate successor of x_1. Then, if X has a smallest element a then clearly the set $U_1 = [a, x_2)$ is a neighborhood (because it is a basis element) of x_1. If X has no smallest element then there is an $a < x_1$ so that $U_1 = (a, x_2)$ is a neighborhood of x_1. Similarly $U_2 = (x_1, b]$ or $U_2 = (x_1, b)$ is a neighborhood of x_2, where b is either the largest element of X or $x_2 < b$, respectively. Either way, for any $y \in U_1$ we have that $y < x_2$ so that $y \leq x_1$ since x_2 is the immediate successor of x_1. Hence it is not true that $y > x_1$ so that $y \notin U_2$. This shows that U_1 and U_2 are disjoint.

Case: x_2 is not the immediate successor of x_1. Then there is an $x \in X$ where $x_1 < x < x_2$. So let $U_1 = [a, x)$ (or $U_1 = (a, x)$) for the smallest element a of X (or some $a < x_1$). Similarly let $U_2 = (x, b]$ (or $U_2 = (x, b)$) for the largest element b of X (or some $x_2 < b$). Either way U_1 and U_2 are neighborhoods of x_1 and x_2, respectively. If $y \in U_1$ then $y < x$ so that clearly it is not true that $y > x$ so that $x \notin U_2$. Hence again U_1 and U_2 are disjoint.

Thus in either case we have shown that X is a Hausdorff space as desired since x_1 and x_2 were an arbitrary pair. $\square$

Exercise 17.11

Show that the product of two Hausdorff spaces is Hausdorff.

Solution:

Proof. Suppose that X and Y are Hausdorff spaces and consider two distinct points (x_1, y_1) and (x_2, y_2) in $X \times Y$. Since these points are distinct, it has to be that $x_1 \neq x_2$ or $y_1 \neq y_2$. In the first case x_1 and x_2 are distinct points of X so that there are disjoint neighborhoods U_1 and U_2 of x_1 and x_2, respectively. This of course follows from the fact that X is a Hausdorff space. Then we have that $U_1 \times Y$ and $U_2 \times Y$ are both basis elements, and therefore open sets, in the product space $X \times Y$ since Y itself is obviously an open set of Y. Clearly also $(x_1, y_1) \in U_1 \times Y$ and $(x_2, y_2) \in U_2 \times Y$ so that $U_1 \times Y$ is a neighborhood of (x_1, y_1) and $U_2 \times Y$ is a neighborhood of (x_2, y_2).

Then, for any $(x, y) \in U_1 \times Y$ we have that $x \in U_1$ so that $x \notin U_2$ since they are disjoint. Then it has to be that $(x, y) \notin U_2 \times Y$. This suffices to show that $U_1 \times Y$ and $U_2 \times Y$ are disjoint since (x, y) was arbitrary. Thus $X \times Y$ is a Hausdorff space since the points (x_1, y_1) and (x_2, y_2) were arbitrary. An analogous argument in the case in which $y_1 \neq y_2$ shows the same result. □

Exercise 17.12

Show that a subspace of a Hausdorff space is Hausdorff.

Solution:

Proof. Suppose that X is a Hausdorff space and that Y is a subset of X. Consider any two distinct points y_1 and y_2 in Y so that of course also $y_1, y_2 \in X$. Then there are neighborhoods U_1 and U_2 of y_1 and y_2, respectively, that are disjoint since X is Hausdorff. Since U_1 is open in X, we have that $V_1 = U_1 \cap Y$ is open in Y by the definition of a subspace topology. Clearly also V_1 contains y_1 since $y_1 \in U_1$ and $y_1 \in Y$. Similarly $V_2 = U_2 \cap Y$ is an open set of Y that contains y_2. Then, for any $x \in V_1$ clearly $x \in U_1$ so that $x \notin U_2$ since U_1 and U_2 are disjoint. Then $x \notin U_2 \cap Y = V_2$. Since x was arbitrary, this shows that V_1 and V_2 are disjoint, which then shows that Y is a Hausdorff space as desired. □

Exercise 17.13

Show that X is Hausdorff if and only if the ***diagonal*** $\Delta = \{x \times x \mid x \in X\}$ is closed in $X \times X$.

Solution:

Proof. ($\Rightarrow$) Suppose that X is Hausdorff and consider any point $x \times y \in X \times X$ where $x \times y \notin \Delta$. Then it must be that $x \neq y$ so that there are disjoint neighborhood U of x and V of y since X is Hausdorff. Then $U \times V$ is a basis element of $X \times X$, by the definition of a product topology, and is therefore open. Now consider any point $w \times z \in U \times V$ so that $w \in U$ and $z \in V$. Then it has to be that $w \neq z$ since U and V are disjoint, which shows that $w \times z \notin \Delta$. Since $w \times z$ was an arbitrary point of $U \times V$, this shows that $U \times V$ does not intersect Δ. Since also $U \times V$ is open and contains $x \times y$, this shows that $x \times y$ is not a limit point of Δ. Moreover, since $x \times y$ was an arbitrary element of $X \times X$ that is not in Δ, it follows that Δ must contain all of its limit points and is therefore closed by Corollary 17.7.

($\Leftarrow$) Now suppose that Δ is closed and suppose that x and y are distinct points in X. Then $x \times y \notin \Delta$ so that $x \times y$ cannot be a limit point of Δ (since it contains all its limit points by Corollary 17.7). Hence there is an open set T in $X \times X$ that contains $x \times y$ and does not intersect Δ. It then follows that there is a basis element $U \times V$ of $X \times X$ containing $x \times y$ where $U \times V \subset T$. Then U and V

are both open in X by the definition of the product topology, and clearly $x \in U$ and $y \in V$. It also follows that $U \times V$ does not intersect Δ since, if it did, then T would as well.

Suppose that U and V are not disjoint so that there is a $z \in U$ where also $z \in V$. Then clearly $z \times z \in U \times V$ but we also have that $z \times z \in \Delta$ so that $U \times V$ intersects Δ. As we know that this cannot be the case, it has to be that U and V are disjoint. This shows that X is Hausdorff as desired since U is a neighborhood of x, V is a neighborhood of y, and x and y were arbitrary distinct points of X. □

Exercise 17.14

In the finite compliment topology on $\mathbb{R}$, to what point or points does the sequence $x_n = 1/n$ converge?

Solution:

We claim that this sequence converges to every point in $\mathbb{R}$.

Proof. Suppose that this is not the case so that there is point $a \in \mathbb{R}$ where the sequence does not converge to A. Then there is an open set U containing a such that, for every $N \in \mathbb{Z}_+$, there is an $n \geq N$ where $x_n \notin U$. It is easy to see that $x_n \notin U$ for an infinite number of $n \in \mathbb{Z}_+$. For, if this were not the case, then there would be an $N \in \mathbb{Z}_+$ where $x_n \in U$ for every $n \geq N$. We know, though, that there must be an $n \geq N$ where $x_n \notin U$.

Moreover, clearly every x_n in the sequence is distinct so that there are an infinite number of points not in U. Since each of these points are still in X, we have that $X - U$ is infinite. As this is the finite compliment topology and U is open, this can only be the case if $X - U = X$ itself, in which case it would have be that $U = \varnothing$ since $U \subset X$. This is not possible since U contains a. So it seems that a contradiction has been reached, which shows the desired result. □

In fact, this is true for any sequence for which the range the sequence $\{x_n \mid n \in \mathbb{Z}_+\}$ is infinite. This is to say that any such sequence converges to every point of $\mathbb{R}$. Note also that this shows that the finite compliment topology on $\mathbb{R}$ is not a Hausdorff space by the contrapositive of Theorem 17.10.

Exercise 17.15

Show the T_1 axiom is equivalent to the condition that for each pair of points of X, each has a neighborhood not containing the other.

Solution:

Note that, though it does not say so above, the points must be distinct since any neighborhood containing x obviously has to contain x.

Proof. ($\Rightarrow$) Suppose that a space X satisfies the T_1 axiom and consider any two distinct points x and y of X. Then the point $\{x\}$ is closed since it is finite, and hence it also contains all of its limit points by Corollary 17.7. Since the point y is not in $\{x\}$ (since $y \neq x$), it cannot be a limit point of $\{x\}$. Hence there is a neighborhood U of y that does not intersect $\{x\}$. Hence $x \notin U$. An analogous argument involving $\{y\}$ shows that there is a neighborhood V of x that does not contain y. Since x and y were arbitrary points, this shows the desired property.

($\Leftarrow$) Now suppose that, for each pair of distinct points in X, each point has a neighborhood that does not contain the other point. As in the proof of Theorem 17.8, it suffices to show that every one-point set is closed, since any finite set can be expressed as the finite union of such sets, which is also then closed by Theorem 17.1. So let $\{x\}$ be such a one-point set and consider any $y \notin \{x\}$ so that clearly $y \neq x$. Then, since x and y are distinct, there is a neighborhood U of y such that U does not contain x. Therefore U and $\{x\}$ are disjoint. This shows that y is not a limit point of $\{x\}$, which shows that $\{x\}$ contains all its limit points since $y \notin \{x\}$ was arbitrary. Hence $\{x\}$ is closed as desired by Corollary 17.7. □

Exercise 17.16

Consider the five topologies on $\mathbb{R}$ given in Exercise 7 of §13.

(a) Determine the closure of the set $K = \{1/n \mid n \in \mathbb{Z}_+\}$ under each of these topologies.

(b) Which of these topologies satisfy the Hausdorff axiom? the T_1 axiom?

Solution:

Lemma 17.16.1. *Suppose that $\mathcal{T}$ and $\mathcal{T}'$ are topologies on X and $\mathcal{T}'$ is finer than $\mathcal{T}$. If $\mathcal{T}$ satisfies the T_1 axiom, then so does $\mathcal{T}'$. Similarly, if $\mathcal{T}$ is Hausdorff, then so is $\mathcal{T}'$.*

Proof. First, suppose that $\mathcal{T}$ satisfies the T_1 axiom and consider any finite subset A of X. Then A is closed in $\mathcal{T}$ by the T_1 axiom so that by definition $X - A$ is open in $\mathcal{T}$ and hence $X - A \in \mathcal{T}$. Then $X - A \in \mathcal{T}'$ as well since $\mathcal{T} \subset \mathcal{T}'$ so that $X - A$ is open in $\mathcal{T}'$. Hence A is closed in $\mathcal{T}'$ by definition. Since A was an arbitrary finite set, this shows that $\mathcal{T}'$ also satisfies the T_1 axiom.

Now suppose that $\mathcal{T}$ is Hausdorff, and consider any two distinct points x and y in X. Then there are neighborhoods U of x and V of y, both in $\mathcal{T}$, that do not intersect since $\mathcal{T}$ is Hausdorff. Then clearly $U, V \in \mathcal{T}'$ as well since $\mathcal{T} \subset \mathcal{T}'$. Hence U and V are neighborhoods of x and y, respectively, in $\mathcal{T}'$ that do not intersect. This shows that $\mathcal{T}'$ is Hausdorff as desired since x and y were arbitrary points of X. □

Main Problem.

First we summarize what we claim about these topologies on $\mathbb{R}$ for both parts:

Topology	Definition	T_1	Hausdorff	$\overline{K}$
$\mathcal{T}_1$	Standard	Yes	Yes	$K \cup \{0\}$
$\mathcal{T}_2$	$\mathbb{R}_K$	Yes	Yes	K
$\mathcal{T}_3$	Finite complement	Yes	No	$\mathbb{R}$
$\mathcal{T}_4$	Upper limit	Yes	Yes	K
$\mathcal{T}_5$	Basis of $(-\infty, a)$ sets	No	No	$\{x \in \mathbb{R} \mid 0 \leq x\}$

Next we justify these claims for each part.

(a) First we show that $\overline{K} = K \cup \{0\}$ in $\mathcal{T}_1$.

Proof. ($\subset$) Consider any real number x and suppose that $x \notin K \cup \{0\}$, hence $x \notin K$ and $x \neq 0$. Since $x \neq 0$, we have

Case: $x < 0$. Then clearly the open set $(x - 1, 0)$ contains x but does not intersect K since $0 < y$ for every $y \in K$, but $y < 0$ for every $y \in (x - 1, 0)$.

Case: $x > 0$. If $1 < x$, then $(1, x + 1)$ contains x but does not intersect K since $y \leq 1$ for every $y \in K$, but $1 < y$ for every $y \in (1, x + 1)$. If $1 \geq x$ it follows from the fact that $x \notin K$ that there is a positive integer n where $n < 1/x < n + 1$, and hence $1/(n + 1) < x < 1/n$. Then clearly the open set $(1/(n+1), 1/n)$ contains x, but we also have that it does not intersect K. If it did, then there would be an integer m where $1/(n+1) < 1/m < 1/n$ so that $n < m < n+1$, which we know is not possible since $n+1$ is the immediate successor of n in $\mathbb{Z}_+$.

Thus in all cases there is a neighborhood of x that does not intersect K. This of course shows that $x \notin \overline{K}$ by Theorem 17.5a. We have therefore shown that $x \notin K \cup \{0\}$ implies that $x \notin \overline{K}$. By contrapositive, this shows that $\overline{K} \subset K \cup \{0\}$.

($\supset$) Now consider any neighborhood U of 0 so that there is a basis element (a, b) containing 0 that is a subset of U. Then $a < 0 < b$. Clearly there is an $n \in \mathbb{Z}_+$ large enough where $a < 0 < 1/n < b$ and hence $1/n \in (a, b) \subset U$. Since also $1/n \in K$, we have that U intersects K. Since U was an arbitrary neighborhood, this shows that 0 is in $\overline{K}$ by Theorem 17.5a. Since also clearly any $x \in K$ is also in $\overline{K}$, it follows that $\overline{K} \supset K \cup \{0\}$. □

Next we show that $\overline{K} = K$ in $\mathcal{T}_2$, which is to say that K is already closed.

Proof. First, clearly $K \subset \overline{K}$ basically by definition. Now consider any $x \notin K$. Then clearly the set $B = (x-1, x+1) - K$ is a basis element of $\mathcal{T}_2$. Also it clearly contains x since $x \notin K$ and also does not intersect K since $y \in B$ means that $y \notin K$. This shows that x is not in $\overline{K}$ by Theorem 17.5b. Since x was arbitrary this shows that $x \notin K$ implies that $x \notin \overline{K}$. Thus $\overline{K} \subset K$ by contrapositive. This suffices to show that $\overline{K} = K$ as desired. □

Now we show that $\overline{K} = \mathbb{R}$ in $\mathcal{T}_3$.

Proof. Consider any real x and any neighborhood U of x. Then U is open in $\mathcal{T}_3$ so that $\mathbb{R} - U$ must be finite, noting that $\mathbb{R} - U$ cannot be all of $\mathbb{R}$ since U would then have to be empty since $U \subset \mathbb{R}$, whereas we know that $x \in U$. It then follows that there are a finite number of real numbers not in U. However, clearly K is an infinite set so that there must be an element of K that *is* in U. This shows that K intersects U so that x is in $\overline{K}$ by Theorem 17.5a since U was an arbitrary neighborhood. Hence $\mathbb{R} \subset \overline{K}$ since x was arbitrary. Clearly also $\overline{K} \subset \mathbb{R}$ so that $\overline{K} = \mathbb{R}$. □

Next we show that $\overline{K} = K$ in $\mathcal{T}_4$ so that K is closed.

Proof. Clearly $K \subset \overline{K}$. So consider any real x where $x \notin K$.

Case: $x \leq 0$. Then the set $B = (x-1, x]$ is clearly a basis element of $\mathcal{T}_4$ that contains x. For any $y \in K$ we have that $x \leq 0 < y$ so that $y \notin B$. Hence B does not intersect K.

Case: $x > 0$. If $1 \leq x$ then it has to be that $1 < x$ since $1 = 1/1 \in K$ but $x \notin K$, and hence $x \neq 1$. Then $B = (1, x]$ is clearly a basis element of $\mathcal{T}_4$ and contains x. This also clearly does not intersect K since $y \leq 1$ for any $y \in K$ so that $y \notin B$. On the other hand, if $1 > x$ then there is an integer n where $n < 1/x < n+1$ so that $1/(n+1) < x < 1/n$ since $x \notin K$. It then follows that the set $B = (1/(n+1), x]$ is a basis element of $\mathcal{T}_4$ that contains x and does not intersect K.

Hence in all cases there is a basis element B containing x that does not intersect K. This shows that $x \notin \overline{K}$ by Theorem 17.5b. Hence we have shown that $x \notin K$ implies that $x \notin \overline{K}$, which shows by contrapositive that $\overline{K} \subset K$. Therefore $\overline{K} = K$ as desired. □

Lastly we show that $\overline{K} = \{x \in \mathbb{R} \mid 0 \leq x\}$ in $\mathcal{T}_5$.

Proof. First, let $A = \{x \in \mathbb{R} \mid 0 \leq x\}$ and consider any $x \in A$ and any basis element $B = (-\infty, a)$ containing x. Hence clearly $0 \leq x$ since $x \in A$ and $x < a$ since $x \in B$. Thus $0 \leq x < a$ so that there is an integer n large enough that $0 < 1/n < a$. Then $1/n \in B$ and also clearly $1/n \in K$. Thus B intersects K. Since B was any neighborhood of x it follows from Theorem 17.5b that $x \in \overline{K}$. Hence $A \subset \overline{K}$ since x was arbitrary.

Now suppose that $x \notin A$ so that $x < 0$. Then the set $B = (-\infty, 0)$ is clearly a basis element of $\mathcal{T}_5$ that contains x. Since $0 < y$ for any $y \in K$, it follows that $y \notin B$, and thus B cannot intersect K. Hence by Theorem 17.5b we have that $x \notin \overline{K}$. This shows that $\overline{K} \subset A$ by contrapositive, which completes the proof that $\overline{K} = A$. $\square$

(b) First we show that $\mathcal{T}_1$, $\mathcal{T}_2$, and $\mathcal{T}_4$ are Hausdorff spaces and satisfy the T_1 axiom.

Proof. First consider any two distinct points $x, y \in \mathbb{R}$. Without loss of generality, we can assume that $x < y$. Let $z = (x+y)/2$ so that clearly $x < z < y$. Then obviously the open intervals $(x-1, z)$ and $(z, y+1)$ are disjoint open sets in $\mathcal{T}_1$ that contain x and y, respectively. This shows that $\mathcal{T}_1$ is a Hausdorff space and therefore also satisfies the T_1 axiom by Theorem 17.8.

It then follows that $\mathcal{T}_2$ and $\mathcal{T}_4$ are also both Hausdorff and satisfy the T_1 axiom. This follows from Lemma 17.16.1 since it was shown in Exercise 13.7 that $\mathcal{T}_1 \subset \mathcal{T}_2 \subset \mathcal{T}_4$. $\square$

Next we show that $\mathcal{T}_3$ satisfies the T_1 axiom but is not a Hausdorff space.

Proof. So first consider any finite subset A of $\mathbb{R}$. Let $U = X - A$ so that clearly $A = X - (X - A) = X - U$. Then, since $X - U = A$ is finite, it follows that U is open in $\mathcal{T}_3$ by the definition of the finite complement topology. Hence by definition A is closed in $\mathcal{T}_3$ since $X - A = U$ is open. This shows that $\mathcal{T}_3$ satisfies the T_1 axiom since A was an arbitrary finite subset of $\mathbb{R}$.

To show that $\mathcal{T}_3$ is not Hausdorff, consider any open set U containing 0 and any open set V containing 1. It then has to be that $\mathbb{R} - U$ is finite since it cannot be that $\mathbb{R} - U = \mathbb{R}$ itself since then U would have to be empty (which we know is not the case since $0 \in U$) since $U \subset \mathbb{R}$. Likewise $\mathbb{R} - V$ is also finite. Thus there are a finite number of real numbers that are not in U and a finite number that are not in V. From this it clearly follows that there are a finite number of real numbers x where $x \notin U$ or $x \notin V$. Since we have

$$x \notin U \vee x \notin V \Leftrightarrow \neg(x \in U \wedge x \in V) \Leftrightarrow \neg(x \in U \cap V) \Leftrightarrow x \notin U \cap V\,,$$

it has to be that there are a finite number of reals numbers that are not in $U \cap V$. But since $\mathbb{R}$ is infinite, this means that there are an infinite number of real numbers that *are* in $U \cap V$. Hence $U \cap V \neq \varnothing$, i.e. they intersect. Since U and V were arbitrary neighborhoods, this shows that $\mathcal{T}_3$ is not Hausdorff by the negation of the definition. $\square$

Lastly we prove that $\mathcal{T}_5$ is neither a Hausdorff space nor satisfies the T_1 axiom.

Proof. First consider the distinct real numbers 0 and 1. Consider then any open set V containing 1 so that there is a basis element $B = (-\infty, a)$ that contains 1 and is a subset of U. Clearly we have that $0 \in B$ since $0 < 1 < a$ and hence $0 \in U$ since $B \subset U$. Since U was an arbitrary neighborhood of 1, it follows there is no neighborhood of 1 that does not contain 0. Hence $\mathcal{T}_5$ does not satisfy the T_1 axiom by the negation of Exercise 17.15. It also then follows that $\mathcal{T}_5$ is not a Hausdorff space by the contrapositive of Theorem 17.8. $\square$

Exercise 17.17

Consider the lower limit topology on $\mathbb{R}$ and the topology given by the basis $\mathcal{C}$ of Exercise 8 of §13. Determine the closures of the intervals $A = (0, \sqrt{2})$ and $B = (\sqrt{2}, 3)$ in these two topologies.

Solution:

Recall that $\mathcal{C} = \{[a, b) \mid a < b,\ a \text{ and } b \text{ rational}\}$ from Exercise 13.8, noting that it was shown there that this basis generates a topology different from the lower limit topology. Denote the lower limit topology by $\mathcal{T}_l$, and denote the topology generated by $\mathcal{C}$ by $\mathcal{T}_c$.

Lemma 17.17.1. *The closure of an open interval (a, b) is $[a, b)$ in the lower limit topology on $\mathbb{R}$.*

Proof. First let $A = (a, b)$ and $C = [a, b)$ so that we must show that $\overline{A} = C$.

($\supset$) Consider any $x \in C$.

Case: $x = a$. Consider any basis element $B = [c, d)$ that contains $x = a$ so that $c \leq x = a < d$. Let $e = \min(b, d)$ so that $a < e$ since both $a < d$ and $a < b$. Then of course there is a real y between a and e so that $a < y < e$. Thus we have $c \leq a < y < e \leq d$ so that $y \in B$. Also $a < y < e \leq b$ so that $y \in A$. Hence B intersects A so that $x = a \in \overline{A}$ by Theorem 17.5b since B was an arbitrary basis element.

Case: $x \neq a$. Then it has to be that $x \in (a, b) = A$ so that $x \in \overline{A}$ since obviously $A \subset \overline{A}$.

This shows that $C \subset \overline{A}$ since x was arbitrary.

($\subset$) Now consider any real x where $x \notin C$ so that either $x < a$ or $x \geq b$. If $x < a$ then the basis element $B = [x, a)$ clearly contains x but does not intersect A. If $x \geq b$ then the basis element $B = [b, x + 1)$ contains x and does not intersect A. Either way it follows from Theorem 17.5b that $x \notin \overline{A}$. Since x was arbitrary, the contrapositive shows that $\overline{A} \subset C$. □

Lemma 17.17.2. *The closure of an open interval (a, b) in $\mathcal{T}_c$ is $[a, b)$ if b is rational and $[a, b]$ if b is irrational.*

Proof. Let $A = (a, b)$. Consider any real x, and we shall consider an exhaustive list of cases that will show whether $x \in \overline{A}$ or $x \notin \overline{A}$.

Case: $x < a$. Obviously there is a rational p where $p < x$ since the rationals are unbounded below. Similarly, there is a rational q where $x < q < a$ since the rationals are order-dense in the reals. The set $B = [p, q)$ is then clearly a basis element of $\mathcal{T}_c$ that contains x. It is also trivial to show that B does not intersect A since $q < a$, which shows that $x \notin \overline{A}$ by Theorem 17.5b whether b is rational or not.

Case: $x = a$. Consider any basis element $B = [p, q)$ (where p and q are rational) that contains $x = a$ so that $p \leq x = a < q$. Let $d = \min(b, q)$ so that $a < d$ since both $a < q$ and $a < b$. Then of course there is a real y between a and d so that $a < y < d$. Thus we have $p \leq a < y < d \leq q$ so that $y \in B$. Also $a < y < d \leq b$ so that $y \in A$. Hence B intersects A so that $x = a \in \overline{A}$ by Theorem 17.5b since B was an arbitrary basis element. Note that this is true whether or not b is rational.

Case: $a < x < b$. Then clearly $x \in (a, b) = A$ so that $x \in \overline{A}$ since obviously $A \subset \overline{A}$.

Case: $x = b$.

> Case: b is rational. Then there is another rational q where $q > b$ since the rationals are unbounded above. Then clearly the set $B = [b, q)$ is a basis element of $\mathcal{T}_c$ that contains b. Also clearly B does not intersect A since $y \in A$ implies that $y < b$ and hence $y \notin B$. This shows that $x = b \notin \overline{A}$ by Theorem 17.5b.
>
> Case: b is irrational. Then consider any basis element $B = [p, q)$ containing b so that p and q are rational. Thus $p \leq b < q$, but since p is rational but b is not, it has to be that $p < b < q$. Let $c = \max(p, a)$ so that $c < b$ since both $a < b$ and $p < b$. There is then a real y where $c < y < b$ so that $a \leq c < y < b$ and hence $y \in A$. Also $p \leq c < y < b < q$ so that also $y \in B$. Therefore B and A intersect, which shows that $x = b \in \overline{A}$ by Theorem 17.5b since B was arbitrary.

Case: $x > b$. Then there are clearly rationals p and q where $b < p < x$ and $x < q$. Then clearly the set $B = [p, q)$ is a basis element that contains x and does not intersect A. This of course shows that $x \notin \overline{A}$ by Theorem 17.5b again, noting that this is true regardless of the rationality of b.

These cases taken together show the desired results. $\square$

Main Problem.

First, it follows directly from Lemma 17.17.1 that that $\overline{A} = [0, \sqrt{2})$ and $\overline{B} = [\sqrt{2}, 3)$ in $\mathcal{T}_l$. It is worth noting that $\overline{A}$ and $\overline{B}$ are both basis elements of $\mathcal{T}_l$, which is interesting since they are closures and therefore closed. This of course implies that basis elements in $\mathcal{T}_l$ are both open and closed, which is indeed the case and is easy to see after a little thought.

It also follows directly from Lemma 17.17.2 that $\overline{A} = [0, \sqrt{2}]$ and $\overline{B} = [\sqrt{2}, 3)$ in $\mathcal{T}_c$ since $\sqrt{2}$ is irrational and 3 is rational.

Exercise 17.18

Determine the closures of the following subsets of the ordered square:

$$\begin{aligned} A &= \{(1/n) \times 0 \mid n \in \mathbb{Z}_+\} \ , \\ B &= \left\{(1 - 1/n) \times \tfrac{1}{2} \mid n \in \mathbb{Z}_+\right\} \ , \\ C &= \{x \times 0 \mid 0 < x < 1\} \ , \\ D &= \left\{x \times \tfrac{1}{2} \mid 0 < x < 1\right\} \ , \\ E &= \left\{\tfrac{1}{2} \times y \mid 0 < y < 1\right\} \ . \end{aligned}$$

Solution:

We assume that the ordered square refers to the set $X = [0, 1]^2$ with the dictionary order topology. Denote the dictionary order on X by $\prec$.

Definition 17.18.1. *For a topology on $\mathbb{R}$ and some subset $A \subset \mathbb{R}$, consider a point $x \in \mathbb{R}$. We say that x is a limit point of A from above if every neighborhood containing x also contains a point y where $y \in A$ and $x < y$. Similarly, a point x is a limit point of A from below if every neighborhood containing x also contains a point y where $y \in A$ and $y < x$.*

Note that a point can be a limit point from both below and above.

Lemma 17.18.2. *Suppose that A is a subset of the real interval $[0, 1]$ and that $B = \{x \times b \mid x \in A\}$ for some $b \in [0, 1]$ so that $B \subset X = [0, 1]^2$. Then the point $x \times y$ is a limit point of B in the dictionary order topology on the unit square if and only if either $y = 1$ and x is a limit point of A from above* or *$y = 0$ and x is a limit point of A from below in the order topology on $[0, 1]$.*

Proof. ($\Rightarrow$) We show this by contrapositive. So suppose that $y \neq 1$ or x is not a limit point of A from above *and* that $y \neq 0$ or x is not a limit point from below.

Case: $y \neq 0$ and $y \neq 1$. Clearly then $0 < y < 1$. If $y = b$ then the dictionary order interval $(x \times 0, x \times 1)$ is a basis element that contains $x \times y$ and that does not contain any other points of B, if indeed $x \in A$ so that $x \times y = x \times b$ is in B. If $y < b$ then the dictionary order interval $(x \times 0, x \times b)$ is a basis element with the same properties. Lastly, if $y > b$ then the dictionary order interval $(x \times b, x \times 1)$ is a basis element that contains $x \times y$ but no points of B.

Case: $y = 0$ or $y = 1$. If $y = 0$ then we have

> Case: $x = 0$. Then, if $b = y = 0$, we have that the dictionary order interval $[0 \times 0, 0 \times 1)$ is a basis element containing $x \times y = 0 \times 0$ but no other points of B, if indeed $x = 0 \in A$ so that

$x \times y \in B$. If $b \neq 0$ then $0 < b$ so that the interval $[0 \times 0, 0 \times b)$ is a basis element with the same properties.

Case: $x \neq 0$. Then $0 < x$ and it has to be that x is a not a limit point of A from below. Thus there is an interval (c, d) or $(c, 1]$ (or $[0, d)$ in which case let $c = 0$ in what follows) that contains x but no other points $y \in A$ where $y < x$. If $b = y = 0$ then it is easy to show that $(c \times 1, x \times 1)$ (or $(c \times 1, x \times 1]$ if $x = 1$) is a basis element that contains $x \times y$ but no other points of B, if indeed $x \in A$ so that $x \times y \in B$. If $b \neq y = 0$ then $0 < b$ so that $(c \times 1, x \times b)$ is a basis element with the same property.

If $y = 1$, then an analogous argument shows analogous results.

Thus in all cases and sub-cases it follows that $x \times y$ is not a limit point of B, which shows the desired result by contrapositive.

($\Leftarrow$) Now suppose that either $y = 1$ and x is a limit point of A from above or $y = 0$ and x is a limit point of A from below. In the first case consider any dictionary order interval $C = (a \times c, d \times e)$ that contains $x \times y$. Then it has to be that $x < d$ since otherwise it would have to be that $y = 1 < e$ since $x \times y \prec d \times e$, which is of course impossible. Then, since x is a limit point of A from above, it follows that the open set $[0, d)$ contains a point $z \in A$ where $x < z$ so that $x < z < d$. It then follows that the point $z \times b$ is in both C and B, and is of course distinct from $x \times y$ since $x < z$. The same argument can be made if C is a basis element in the form of $[0 \times 0, d \times e)$ or $(a \times c, 1 \times 1]$. This suffices to show that $x \times y$ is a limit point of B since C was an arbitrary basis element.

An analogous argument can be made in the case when $y = 0$ and x is a limit point of A from below, which shows the desired result. $\square$

Main Problem.

First we claim that $\overline{A} = A \cup \{0 \times 1\}$.

Proof. First, let $K = \{1/n \mid n \in \mathbb{Z}_+\} \subset [0, 1]$ so that clearly $A = \{x \times 0 \mid x \in K\}$. It is easy to show that 0 is the only limit point of K and it is a limit point from above only. It then follows from Lemma 17.18.2 that 0×1 is the only limit point of A so that $\overline{A} = A \cup \{0 \times 1\}$ since the closure is the union of the set and the set of its limit points. $\square$

Next we claim that $\overline{B} = B \cup \{1 \times 0\}$.

Proof. This time let $L = \{1 - 1/n \mid n \in \mathbb{Z}_+\}$ so that clearly $B = \left\{x \times \frac{1}{2} \mid x \in L\right\}$. It is trivial to show that 1 is the only limit point of L and that it is a limit point from below only. Hence 1×0 is the only limit point of B by Lemma 17.18.2 so that the result follows. $\square$

Now we claim that $\overline{C} = C \cup \{1 \times 0\} \cup \{x \times 1 \mid 0 \leq x < 1\}$.

Proof. First, we clearly have that $C = \{x \times 0 \mid x \in (0, 1)\}$. It is easy to show that every point of $(0, 1)$ is a limit point both from above and below, that 0 is a limit point from above only, and that 1 is a limit point from below only. Thus it follows that the set of limit points of C are then $\{x \times 0 \mid 0 < x \leq 1\} \cup \{x \times 1 \mid 0 \leq x < 1\}$ by Lemma 17.18.2. As many of these points are contained in C itself, the result follows. $\square$

We claim that $\overline{D} = D \cup \{x \times 0 \mid 0 < x \leq 1\} \cup \{x \times 1 \mid 0 \leq x < 1\}$.

Proof. The limit points of D are the same as for C above for the same reasons, i.e. $\{x \times 0 \mid 0 < x \leq 1\} \cup \{x \times 1 \mid 0 \leq x < 1\}$. The result then follows. $\square$

Lastly, we claim that $\overline{E} = \left\{\frac{1}{2} \times y \mid 0 \leq y \leq 1\right\} = \left\{\frac{1}{2}\right\} \times [0, 1]$, noting that clearly $E = \left\{\frac{1}{2}\right\} \times (0, 1)$.

Proof. Let $F = \left\{\frac{1}{2}\right\} \times [0,1]$ so that we must show that $\overline{E} = F$.

($\subset$) Consider any $x \times y$ where $x \times y \notin F$ so that simply $x \neq \frac{1}{2}$ since it has to be that $y \in [0,1]$. If $x < \frac{1}{2}$ then the basis element $[0 \times 0, \frac{1}{2} \times 0)$ clearly contains $x \times y$ but no elements of E. If $x > \frac{1}{2}$ then the basis element $(\frac{1}{2} \times 1, 1 \times 1]$ clearly contains $x \times y$ but no elements of E either. This shows that $x \times y$ is a not in $\overline{E}$ by Theorem 17.5b. Hence $\overline{E} \subset F$ by contrapositive.

($\supset$) Consider any $x \times y \in F$ so that $x = \frac{1}{2}$ and $y \in [0,1]$. If $y \in (0,1)$ then $x \times y \in E$ so that $x \times y \in \overline{E}$ since obviously $E \subset \overline{E}$. If $y = 0$ then consider any dictionary order interval $F = (a \times c, b \times d)$ containing $x \times y = \frac{1}{2} \times 0$. In particular we have that $\frac{1}{2} \times 0 \prec b \times d$ so that either $\frac{1}{2} < b$, or $b = \frac{1}{2}$ and $0 < d$. In the first case we have that $\frac{1}{2} \times \frac{1}{2}$ is in both F and E. In the second case let $z = d/2$ so that we have $0 < z < d \leq 1$. Then clearly the point $\frac{1}{2} \times z$ is in F, but we also have that $\frac{1}{2} \times z$ is in E since $0 < z < 1$. The same argument applies if the basis element F is of the form $[0 \times 0, b \times d)$ or $(a \times c, 1 \times 1]$. A similar argument shows an analogous result in the case when $y = 1$. This shows by Theorem 17.5b that $x \times y \in \overline{E}$ since F was an arbitrary basis element, which of course shows that $\overline{E} \supset F$ since $x \times y$ was arbitrary. $\square$

Exercise 17.19

If $A \subset X$, we define the ***boundary*** of A by the equation

$$\operatorname{Bd} A = \overline{A} \cap \overline{(X - A)}\,.$$

(a) Show that $\operatorname{Int} A$ and $\operatorname{Bd} A$ are disjoint, and $\overline{A} = \operatorname{Int} A \cup \operatorname{Bd} A$.

(b) Show that $\operatorname{Bd} A = \varnothing \Leftrightarrow A$ is both open and closed.

(c) Show that U is open $\Leftrightarrow \operatorname{Bd} U = \overline{U} - U$.

(d) If U is open, is it true that $U = \operatorname{Int}(\overline{U})$? Justify your answer.

Solution:

(a)

Proof. Consider any $x \in \operatorname{Int} A$ so that there is a neighborhood of x that is entirely contained in A. Then, for any $y \in U$, we have that $y \in A$ and hence $y \notin X - A$. This shows that U does not intersect $X - A$, which suffices to show that x is not in the closure of $X - A$ by Theorem 17.5a. Thus x is not in the boundary of A since $\operatorname{Bd} A = \overline{A} \cap \overline{(X - A)}$. This of course shows that $\operatorname{Int} A$ and $\operatorname{Bd} A$ are disjoint since x was arbitrary.

To show that $\overline{A} = \operatorname{Int} A \cup \operatorname{Bd} A$, first consider any $x \in \overline{A}$. If $x \in \operatorname{Int} A$ then clearly $x \in \operatorname{Int} A \cup \operatorname{Bd} A$, so assume that $x \notin \operatorname{Int} A$. Consider any neighborhood U of X. Then it has to be that U is not a subset of A since otherwise x would be in the union of open subsets of A and hence in the interior. It then follows that there is a point $y \in U$ where $y \notin A$ and therefore $y \in X - A$. This shows that U intersects $X - A$ so that x is in the closure of $X - A$ since U was an arbitrary neighborhood. Since also $x \in \overline{A}$, we have that $x \in \overline{A} \cap \overline{(X - A)} = \operatorname{Bd} A$. Hence clearly $x \in \operatorname{Int} A \cup \operatorname{Bd} A$ so that $\overline{A} \subset \operatorname{Int} A \cup \operatorname{Bd} A$ since x was arbitrary.

Now consider any $x \in \operatorname{Int} A \cup \operatorname{Bd} A$. If $x \in \operatorname{Int} A$ then also $x \in \overline{A}$ since we have that $\operatorname{Int} A \subset A \subset \overline{A}$. On the other hand, if $x \in \operatorname{Bd} A = \overline{A} \cap \overline{(X - A)}$ then of course $x \in \overline{A}$. This shows that $\operatorname{Int} A \cup \operatorname{Bd} A \subset \overline{A}$ in either case since x was arbitrary. Since both directions have been shown, it follows that $\overline{A} = \operatorname{Int} A \cup \operatorname{Bd} A$ as desired. $\square$

(b)

Proof. ($\Rightarrow$) First suppose that $\operatorname{Bd} A = \varnothing$. Then by part (a) we have that $\overline{A} = \operatorname{Int} A \cup \operatorname{Bd} A = \operatorname{Int} A \cup \varnothing = \operatorname{Int} A$. Hence $A \subset \overline{A} = \operatorname{Int} A$ so that $A = \operatorname{Int} A$ since it is also always the case that $\operatorname{Int} A \subset A$. This shows that A is open since $\operatorname{Int} A$ is always open. We also have $\overline{A} = \operatorname{Int} A \subset A$ so that $A = \overline{A}$ since it is always also the case that $A \subset \overline{A}$. This of course shows that A is also closed since $\overline{A}$ is always closed.

($\Leftarrow$) Now suppose that A is both open and closed. It then follows that $\overline{A} = A = \operatorname{Int} A$. So consider any $x \in \overline{A}$ so that also $x \in \operatorname{Int} A$. Then there is a neighborhood U of x contained entirely in A. Thus, for any point $y \in U$, we have that $y \in A$ so that $y \notin X - A$, which shows that U does not intersect $X - A$. Since U is a neighborhood of x, this shows that $x \notin \overline{X - A}$ by Theorem 17.5a. Then, since x was an arbitrary element of $\overline{A}$, it follows that $\overline{A}$ and $\overline{X - A}$ are disjoint so that $\operatorname{Bd} A = \overline{A} \cap \overline{(X - A)} = \varnothing$ as desired. $\square$

(c)

Proof. ($\Rightarrow$) First suppose that U is open and consider any $x \in \operatorname{Bd} U$. Then we have that $x \in \overline{U}$ and $x \in \overline{X - U}$ since $\operatorname{Bd} U = \overline{U} \cap \overline{(X - U)}$ by definition. Suppose for the moment that $x \in U$ so that U itself is a neighborhood of x since it is open. For any $y \in U$ we have that $y \notin X - U$, and hence U does not intersect $X - U$. This shows that x is not in $\overline{X - U}$ by Theorem 17.5a, which is a contradiction since we know it is. Thus it must be that $x \notin U$ so that $x \in \overline{U} - U$. This of course shows that $\operatorname{Bd} U \subset \overline{U} - U$ since x was arbitrary.

Now consider any $x \in \overline{U} - U$ so that clearly $x \in \overline{U}$. Since also $x \notin U$, it follows that $x \in X - U$ so that of course $x \in \overline{X - U}$ as well. Hence $x \in \overline{U} \cap \overline{(X - U)} = \operatorname{Bd} U$, which shows that $\overline{U} - U \subset \operatorname{Bd} U$ since x was arbitrary. This suffices to show that $\operatorname{Bd} U = \overline{U} - U$ as desired.

($\Leftarrow$) Now suppose that $\operatorname{Bd} U = \overline{U} - U$ and consider any $x \in U$. Then we have that $x \notin \overline{U} - U = \operatorname{Bd} U = \overline{U} \cap \overline{(X - U)}$. Since we know that $x \in \overline{U}$ (since $U \subset \overline{U}$), it must be that $x \notin \overline{X - U}$. Thus, by Theorem 17.5a, there is a neighborhood of V of x that does not intersect $X - U$. This means that, for any point $y \in V$, we have that $y \notin X - U$. Since of course $y \in X$, it follows that y must be in U. This shows that $V \subset U$ since y was arbitrary. Hence V is a neighborhood of x that is entirely contained in U so that x is in the union of open sets contained in U, hence $x \in \operatorname{Int} U$. Since x was an arbitrary element of U, this shows that $U \subset \operatorname{Int} U$. As it is always the case that $\operatorname{Int} U \subset U$ as well, we have that $U = \operatorname{Int} U$ so that U is open since $\operatorname{Int} U$ is always open. $\square$

(d) We claim that this is not generally true.

Proof. As a counterexample consider the set $U = \mathbb{R} - \{0\}$ in the finite complement topology on $\mathbb{R}$. Clearly U is open as its complement $\mathbb{R} - U = \{0\}$ is finite. It is also obvious that U is an infinite set.

Now consider any real number x and any neighborhood V of x. It cannot be that $\mathbb{R} - V$ is all of $\mathbb{R}$ since then V would be empty, and we know that $x \in V$. So it must be that $\mathbb{R} - V$ is finite since V is open, which means that there are only a finite number of real numbers that are *not* in V. However, since U is infinite, there must be an element of U that *is* in V (in fact there are an infinite number of such elements). Hence V intersects U so that $x \in \overline{U}$ by Theorem 17.5a. Since $x \in \mathbb{R}$ was arbitrary, it must be that $\overline{U}$ is all of $\mathbb{R}$.

Clearly $\mathbb{R}$ is open (since the a set is always open in any topology on that set) so that $\operatorname{Int}(\overline{U}) = \operatorname{Int} \mathbb{R} = \mathbb{R}$. Then, since $0 \in \mathbb{R} = \operatorname{Int}(\overline{U})$ but $0 \notin U$, we have that $U \neq \operatorname{Int}(\overline{U})$. $\square$

Exercise 17.20

Find the boundary and the interior of each of the following subsets of $\mathbb{R}^2$:

(a) $A = \{x \times y \mid y = 0\}$

(b) $B = \{x \times y \mid x > 0 \text{ and } y \neq 0\}$
(c) $C = A \cup B$
(d) $D = \{x \times y \mid x \text{ is rational}\}$
(e) $E = \left\{x \times y \mid 0 < x^2 - y^2 \leq 1\right\}$
(f) $F = \{x \times y \mid x \neq 0 \text{ and } y \leq 1/x\}$

Solution:

(a) It is easy to show that A is closed so that $\overline{A} = A$, and that also $\overline{\mathbb{R} - A} = A$ so that $\operatorname{Bd} A = A$. It is also easy to see that no basis element and therefore no neighborhood of any point in A is contained entirely within A. From this it follows that $\operatorname{Int} A = \varnothing$.

(b) It is easy to show that B is open so that $\operatorname{Int} B = B$. It is likewise not difficult to prove that $\overline{B} = \{x \times y \mid x \geq 0\}$. We then have from Exercise 17.19c that $\operatorname{Bd} B = \overline{B} - B = \{x \times y \mid x = 0\} \cup \{x \times y \mid x > 0 \text{ and } y = 0\}$.

(c) Here we have that $C = A \cup B = \{x \times y \mid y = 0\} \cup \{x \times y \mid x > 0\}$. It is then easy to show that the closure is $\overline{C} = \{x \times y \mid y = 0\} \cup \{x \times y \mid x \geq 0\}$. We also have that $\mathbb{R} - C = \{x \times y \mid x \leq 0 \text{ and } y \neq 0\}$ so that $\overline{\mathbb{R} - C} = \{x \times y \mid x \leq 0\}$. From these we clearly then have

$$\operatorname{Bd} C = \overline{C} \cap \overline{(\mathbb{R} - C)} = \{x \times y \mid x < 0 \text{ and } y = 0\} \cup \{x \times y \mid x = 0\} .$$

It is also not difficult to show that $\operatorname{Int} C = \{x \times y \mid x > 0\}$.

(d) Clearly we have that $\overline{D}$ is all of $\mathbb{R}^2$ as a consequence of the fact that the rationals are order-dense in the reals. Also, since any neighborhood of any point in D will intersect a point $x \times y$ with irrational x, it follows that no point of D is in its interior. Thus $\operatorname{Int} D = \varnothing$ so that $\overline{D} = \operatorname{Int} D \cup \operatorname{Bd} D = \varnothing \cup \operatorname{Bd} D = \operatorname{Bd} D$ by Exercise 17.19a, and hence $\operatorname{Bd} D = \overline{D} = \mathbb{R}^2$.

(e) It should be fairly obvious by this point that

$$\operatorname{Bd} E = \{x \times y \mid |y| = |x|\} \cup \left\{x \times y \mid x^2 - y^2 = 1\right\}$$

and $\operatorname{Int} E = \left\{x \times y \mid 0 < x^2 - y^2 < 1\right\}$. This would be easy but tedious to prove rigorously.

(f) First we clearly have that $\operatorname{Int} F = \{x \times y \mid x \neq 0 \text{ and } y < 1/x\}$. We also have that $\overline{F} = \{x \times y \mid x = 0\} \cup \{x \times y \mid x \neq 0 \text{ and } y \leq 1/x\}$. By Exercise 17.19a we have that $\overline{F} = \operatorname{Int} F \cup \operatorname{Bd} F$ and that $\operatorname{Int} F \cap \operatorname{Bd} F = \varnothing$ so that $\operatorname{Bd} F = \overline{F} - \operatorname{Int} F$. Thus we have that $\operatorname{Bd} F = \{x \times y \mid x = 0\} \cup \{x \times y \mid x \neq 0 \text{ and } y = 1/x\}$. Again these facts are not difficult to show rigorously but would be tedious.

Exercise 17.21

(Kuratowski) Consider the collection of all subsets A of the topological space X. The operations of closure $A \to \overline{A}$ and complementation $A \to X - A$ are functions from the collection to itself.

(a) Show that starting with a given set A, one can form no more than 14 distinct sets by applying these two operations successively.
(b) Find a subset A of $\mathbb{R}$ (in its usual topology) for which the maximum of 14 is obtained.

Solution:

For the following we introduce the following notation to make things simpler. If A is a subset of a topological space X then denote

$$cA = \overline{A} \qquad xA = X - C$$
$$iA = \operatorname{Int} A \qquad bA = \operatorname{Bd} A\,.$$

We can consider these (c, x, i, and b) as operators on sets that can be chained together in the obvious way so that, for example, $cxiA = \overline{X - \operatorname{Int} A}$.

Lemma 17.21.1. *For a subset A of topological space X, $X = cA \cup ixA$, and cA and ixA are disjoint*

Proof. First, it is obvious that $cA \cup ixA \subset X$ since each of the sets in the union is a subset of X. Now consider any $x \in X$ and suppose that $x \notin cA = \overline{A}$. Then by Lemma 17.5a there is an open set U containing x where U does not intersect A. For any $y \in U$ we thus have that $y \notin A$ and hence $y \in X - A = xA$. This shows that $U \subset xA$ since y was arbitrary, which suffices to show that $x \in \operatorname{Int}(xA) = ixA$ since U is a neighborhood of x. This of course shows that $x \in cA \cup ixA$ so that $X \subset cA \cup ixA$ since x was arbitrary. This completes the proof that $X = cA \cup ixA$.

To show that cA and ixA are disjoint, consider any $x \in cA$. Consider any neighborhood U of x so that U intersects A by Lemma 17.5a. Hence there is a point $y \in U$ where also $y \in A$, from which it follows that $y \notin X - A = xA$. This suffices to show that U is not a subset of xA. Since U is an arbitrary neighborhood, this shows that $x \notin \operatorname{Int}(xA) = ixA$. This of course shows that cA and ixA are disjoint as desired. □

Lemma 17.21.2. *For a subsets A and B of topological space X where $A \subset B$, we have the following:*

(a) $cA \subset cB$

(b) $iA \subset iB$

(c) $ccA = cA$

(d) $iiA = iA$

(e) $xxA = A$

(f) $xcA = ixA$

(g) $xiA = cxA$

(h) $icicA = icA$

(i) $ciciA = ciA$.

Proof. (a) This was shown in Exercise 17.6a.

(b) Consider any $x \in iA$ so that there is a neighborhood U of x that is totally contained in A. Then clearly U is also totally contained in B as well since, for any $x \in U$, we have that $x \in A$ and hence $x \in B$ since $A \subset B$. This shows that $x \in iB$ since U is a neighborhood of x. Hence $iA \subset iB$ since x was arbitrary.

(c) Since $cA = \overline{A}$ is closed, we clearly have $ccA = cA$.

(d) Since $iA = \operatorname{Int} A$ is open, its interior is itself, i.e. $iiA = iA$.

(e) Obviously $xxA = X - (X - A) = A$ since $A \subset X$.

(f) We have by Lemma 17.21.1 that $X = cA \cup ixA$ where cA, and ixA are mutually disjoint. From this it follows that $ixA = X - cA = xcA$.

(g) We have

$$\begin{aligned} cxA &= xxcxA && \text{(by (e))} \\ &= xixxA && \text{(by (f))} \\ &= xiA && \text{(by (e) again)} \end{aligned}$$

as desired.

(h) First we have that $icA = iicA$ by (d). Also clearly $icA = c(icA) = cicA$ since a set is always a subset of its closure. Hence by (b) we have that $icA = iicA = i(icA) \subset i(cicA) = icicA$. Now, we also have that $icA = i(cA) \subset cA$ since the interior of a set is always a subset of the set. Hence by (a) and (c) we have $cicA = c(icA) \subset c(cA) = ccA = cA$. It then follows from (b) that $icicA = i(cicA) \subset i(cA) = icA$ as well. This of course shows that $icicA = icA$ as desired.

(i) Lastly, we have

$$\begin{aligned}
ciciA &= cicixxA && \text{(by (e))} \\
&= cicxcxA && \text{(by (f))} \\
&= cixicxA && \text{(by (g))} \\
&= cxcicxA && \text{(by (f))} \\
&= xicicxA && \text{(by (g))} \\
&= xicxA && \text{(by (h))} \\
&= cxcxA && \text{(by (g))} \\
&= cixxA && \text{(by (f))} \\
&= ciA && \text{(by (e))}
\end{aligned}$$

as desired. □

Main Problem.

(a)

Proof. We are interested in sequences applying the operators c and x to a subset A. By Lemma 17.21.2 (c) and (e) we have that $ccA = cA$ and $xxA = A$. Thus there is no point in ever applying c or x twice in a row since that would clearly result in a set that we have seen before. We are then interested only in sequences that apply alternating c and x. If we apply the closure c first, we obtain the following sequence:

$$\begin{aligned}
A &= A \\
cA &= cA \\
xcA &= ixA && \text{(by Lemma 17.21.2f)} \\
cxcA &= cixA && \text{(previous result)} \\
xcxcA &= xcixA && \text{(previous result)} \\
&= ixixA && \text{(by Lemma 17.21.2f)} \\
&= icxxA && \text{(by Lemma 17.21.2g)} \\
&= icA && \text{(by Lemma 17.21.2e)} \\
cxcxcA &= cicA && \text{(previous result)} \\
xcxcxcA &= xcicA && \text{(previous result)} \\
&= ixicA && \text{(by Lemma 17.21.2f)} \\
&= icxcA && \text{(by Lemma 17.21.2g)} \\
&= icixA && \text{(by Lemma 17.21.2f)}
\end{aligned}$$

If we apply the next operation we obtain

$$\begin{aligned}
cxcxcxcA &= cicixA && \text{(previous result)} \\
&= cixA\,, && \text{(by Lemma 17.21.2i)}
\end{aligned}$$

which is the same as the fourth set above. Therefore we can get at most 7 distinct sets by applying c first, including A itself. If we instead apply x first then we get the following sequence:

$$xA = xA$$

$$\begin{aligned}
cxA &= cxA \\
xcxA &= ixxA && \text{(corresponding result above)} \\
&= iA && \text{(by Lemma 17.21.2e)} \\
cxcxA &= ciA && \text{(previous result)} \\
xcxcxA &= xciA && \text{(previous result)} \\
&= ixiA && \text{(by Lemma 17.21.2f)} \\
&= icxA && \text{(by Lemma 17.21.2g)} \\
cxcxcxA &= cicxA && \text{(previous result)} \\
xcxcxcxA &= xcicxA && \text{(previous result)} \\
&= ixicxA && \text{(by Lemma 17.21.2f)} \\
&= icxcxA && \text{(by Lemma 17.21.2g)} \\
&= icixxA && \text{(by Lemma 17.21.2f)} \\
&= iciA && \text{(by Lemma 17.21.2e)}
\end{aligned}$$

Again if we try to apply the next operation we get

$$\begin{aligned}
cxcxcxcxA &= ciciA && \text{(previous result)} \\
&= ciA && \text{(by Lemma 17.21.2i)}
\end{aligned}$$

which as before is the same as the fourth set in the sequence. Hence we have at most 7 distinct sets in this sequence for a total of 14 potentially distinct sets as desired. □

Note that this only shows that there can be *no more than* 14 distinct sets. It could be that there are always less than 14 in general. While there are certainly sets that generate less than 14 distinct sets, the next part shows the existence of a topology and a set that does result in 14 distinct sets. This of course shows that 14 is the lowest possible bound in general.

(b) We claim that $A = (-3,-2) \cup (-2,-1) \cup ([0,1] \cap \mathbb{Q}) \cup \{2\}$ in the standard topology on $\mathbb{R}$ is a set that results in 14 distinct sets when the operational sequences from part (a) are applied. We do not prove each sequential operation as this is easy but would be prohibitively tedious. First we enumerate the first sequence, starting with A.

Operations	Set
A	$(-3,-2) \cup (-2,-1) \cup ([0,1] \cap \mathbb{Q}) \cup \{2\}$
cA	$[-3,-1] \cup [0,1] \cup \{2\}$
$xcA = ixA$	$(-\infty,-3) \cup (-1,0) \cup (1,2) \cup (2,\infty)$
$cxcA = cixA$	$(-\infty,-3] \cup [-1,0] \cup [1,\infty)$
$xcxcA = icA$	$(-3,-1) \cup (0,1)$
$cxcxcA = cicA$	$[-3,-1] \cup [0,1]$
$xcxcxcA = icixA$	$(-\infty,-3) \cup (-1,0) \cup (1,\infty)$

Next we enumerate the next sequence of 7 sets, starting with xA:

Operations	Set
xA	$(-\infty,-3] \cup \{-2\} \cup [-1,0) \cup ((0,1) - \mathbb{Q}) \cup (1,2) \cup (2,\infty)$
cxA	$(-\infty - 3] \cup \{-2\} \cup [-1,\infty)$
$xcxA = iA$	$(-3,-2) \cup (-2,-1)$
$cxcxA = ciA$	$[-3,-1]$
$xcxcxA = icxA$	$(-\infty,-3) \cup (-1,\infty)$
$cxcxcxA = cicxA$	$(-\infty,-3] \cup [-1,\infty)$
$xcxcxcxA = iciA$	$(-3,1)$

It is easy to see that these are 14 distinct sets.

We do note that, in an interval containing only rationals (or only irrationals), such as $[0,1] \cap \mathbb{Q}$ used as part of A, clearly every point in the interval is a limit point, including any irrational (or rational) points. This is because any open interval containing any real always contains both rationals and irrationals on account of $\mathbb{Q}$ being order-dense in $\mathbb{R}$. For the same reason no point of such an interval of rationals (or irrationals) is in its interior. If, for example $C = [0,1] \cap \mathbb{Q}$, this clearly then results in $cC = [0,1]$ and $iC = \varnothing$. Indeed this property of this part of A is crucial in its success in generating 14 distinct sets.

§18 Continuous Functions

Exercise 18.1

Prove that for functions $f : \mathbb{R} \to \mathbb{R}$, the ϵ-δ definition of continuity implies the open set definition.

Solution:

Recall that that $f : \mathbb{R} \to \mathbb{R}$ is continuous at a point $x \in \mathbb{R}$ if, for every real $\epsilon > 0$, there is a real $\delta > 0$ such that $|f(y) - f(x)| < \epsilon$ for every real y where $|y - x| < \delta$. We say that f itself is continuous if it is continuous at every $p \in \mathbb{R}$.

Proof. Suppose that $f : \mathbb{R} \to \mathbb{R}$ is continuous by the ϵ-δ definition above. We show that this implies the open set definition by showing that f satisfies (4) in Theorem 18.1. So consider any $x \in \mathbb{R}$ and any neighborhood V of $f(x)$. Then of course there is a basis element (c,d) containing $f(x)$ such that $(c,d) \subset V$. Let $\epsilon = \min(f(x) - c, d - f(x))$, noting that $\epsilon > 0$ since $c < f(x) < d$. It is then trivial to show that $(f(x) - \epsilon, f(x) + \epsilon) \subset (c,d) \subset V$ and contains x.

Then, since f is continuous at x, there is $\delta > 0$ such that $|y - x| < \delta$ implies that $|f(y) - f(x)| < \epsilon$ for any real y. Let $U = (x-\delta, x+\delta)$, which is clearly a neighborhood of x. Now consider any $z \in f(U)$ so that $z = f(y)$ for some $y \in U$. Then we have that $x-\delta < y < x+\delta$ so that clearly $-\delta < y-x < \delta$, from which it follows that $|y - x| < \delta$. We then know that $|z - f(x)| = |f(y) - f(x)| < \epsilon$ since f is continuous. Hence $-\epsilon < z - f(x) < \epsilon$ so that $f(x) - \epsilon < z < f(x) + \epsilon$, and thus $z \in V$ since $(f(x) - \epsilon, f(x) + \epsilon) \subset V$. Since $z \in f(U)$ was arbitrary, this shows that $f(U) \subset V$, which shows that (4) holds for f since x was also arbitrary. □

Exercise 18.2

Suppose that $f : X \to Y$ is continuous. If x is a limit point of the subset A of X, is it necessarily true the $f(x)$ is a limit point of $f(A)$?

Solution:

This is not necessarily true.

Proof. As a counterexample consider a constant function $f : X \to Y$ defined by $f(x) = y_0$ for any $x \in X$ and some $y_0 \in Y$. It was shown in Theorem 18.2a that this is continuous. However, clearly $f(A) = \{y_0\}$ for any subset A of X. So even if x is a limit point of A, no neighborhood of $f(x)$ can intersect $f(A)$ in a point other than $f(x) = y_0$ since y_0 is the only point in $f(A)$! Therefore $f(x)$ is not a limit point of $f(A)$. □

Exercise 18.3

Let X and X' denote a single set in two topologies $\mathcal{T}$ and $\mathcal{T}'$, respectively. Let $i : X' \to X$ be the identity function.

(a) Show that i is continuous $\Leftrightarrow$ $\mathcal{T}'$ is finer than $\mathcal{T}$.
(b) Show that i is a homeomorphism $\Leftrightarrow$ $\mathcal{T}' = \mathcal{T}$.

Solution:

(a)

Proof. First note that clearly the inverse of the identity function is itself with the domain and range reversed, and that for any subset $A \subset X = X'$ we have $i(A) = i^{-1}(A) = A$.

($\Rightarrow$) Suppose that i is continuous and consider any open set $U \in \mathcal{T}$. Then we have that $i^{-1}(U) = U$ is open in $\mathcal{T}'$ since i is continuous. Since U was arbitrary, this shows that $\mathcal{T} \subset \mathcal{T}'$ so that $\mathcal{T}'$ is finer.

($\Leftarrow$) Now suppose that $\mathcal{T}'$ is finer so that $\mathcal{T} \subset \mathcal{T}'$. Consider any open set $U \in \mathcal{T}$ so that also clearly $U \in \mathcal{T}'$, i.e. U is also open in $\mathcal{T}'$. Since $i^{-1}(U) = U$, this shows that i is continuous by the definition of continuity. $\square$

(b)

Proof. Clearly i is a bijection since its domain and range are the same set, and $i^{-1} = i$. We then have that

$$\begin{aligned} i \text{ is a homeomorphism } &\Leftrightarrow i \text{ and } i^{-1} \text{ are both continuous} \\ &\Leftrightarrow \mathcal{T}' \text{ is finer than } \mathcal{T} \text{ and } \mathcal{T} \text{ is finer than } \mathcal{T}' \quad \text{(by part (a) applied twice)} \\ &\Leftrightarrow \mathcal{T} \subset \mathcal{T}' \text{ and } \mathcal{T}' \subset \mathcal{T} \\ &\Leftrightarrow \mathcal{T}' = \mathcal{T} \end{aligned}$$

as desired. $\square$

Exercise 18.4

Given $x_0 \in X$ and $y_0 \in Y$, show that the maps $f : X \to X \times Y$ and $g : Y \to X \times Y$ defined by

$$f(x) = x \times y_0 \qquad \text{and} \qquad g(y) = x_0 \times y$$

are imbeddings.

Solution:

We only show that f is an imbedding of X in $X \times Y$ as the argument for g is entirely analogous.

Proof. First, it is easy to see and trivial to formally show that f is injective. The function f can be of course be defined as $f(x) = f_1(x) \times f_2(x)$ where $f_1 : X \to X$ is the identity function and $f_2 : X \to Y$ is the constant function that maps every element of X to y_0. Since these have both been proven to be continuous in the text, it follows that f is continuous by Theorem 18.4.

Now let f' be the function obtained by restricting the range of f to $f(X) = \{x \times y_0 \mid x \in X\}$. Since f is injective, it follows that f' is a bijection. It follows from Theorem 18.2e that f' is continuous. Clearly the inverse function f'^{-1} is equal to the projection function π_1 so that $f'^{-1}(x, y) = x$. This was shown to be continuous in the proof of Theorem 18.4. This suffices to show that f' is a homeomorphism, which shows the f is an imbedding of X in $X \times Y$. $\square$

Exercise 18.5

Show that the subspace (a, b) of $\mathbb{R}$ is homeomorphic with $(0, 1)$ and the subspace $[a, b]$ of $\mathbb{R}$ is homeomorphic with $[0, 1]$.

Solution:

First we show that (a, b) is homeomorphic to $(0, 1)$.

Proof. First let $X = (a, b)$ and $Y = (0, 1)$, and define the map $f : X \to Y$ by

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

for any $x \in X$, noting that this is defined since $a < b$ so that $b - a > 0$. It is trivial to show that f is a bijection.

Now, f is a linear function that could just as well be defined as a map from $\mathbb{R}$ to $\mathbb{R}$, and clearly this would be continuous by basic calculus. It then follows from Theorem 18.2d that restricting its domain to X means that it is still continuous. We also clearly have from basic algebra that its inverse is the function $f^{-1} : Y \to X$ defined by

$$f^{-1}(y) = a + y(b - a)$$

for $y \in Y$. As this is also linear, it too is continuous by the same argument. This suffices to show that f is a homeomorphism. $\square$

The exact same argument shows that $[a, b]$ is homeomorphic to $[0, 1]$ by simply setting $X = [a, b]$ and $Y = [0, 1]$ in the above proof. It is assumed that here again $a < b$ even though the interval $[a, b]$ is valid if $a = b$ and simply becomes $[a, b] = [a, a] = \{a\}$. However, clearly this set cannot be homeomorphic to $[0, 1]$ since it is finite whereas $[0, 1]$ is uncountable.

Exercise 18.6

Find a function $f : \mathbb{R} \to \mathbb{R}$ that is continuous at precisely one point.

Solution:

For any real x define

$$f(x) = \begin{cases} 0 & x \in \mathbb{Q} \\ x & x \notin \mathbb{Q}. \end{cases}$$

We claim that this is continuous only at $x = 0$.

Proof. As it is easier to do so, we show this using the ϵ-δ definition of continuity, which we know implies the open set definition by Exercise 18.1. First we note that $f(0) = 0$ since 0 is rational. Now consider any $\epsilon > 0$ and let $\delta = \epsilon$. Suppose real y where $|y - 0| = |y| < \delta$. If y is rational then $y = 0$ so that $|y - f(0)| = |0 - 0| = |0| = 0 < \epsilon$. If y is irrational then $|y - f(0)| = |y - 0| = |y| < \delta = \epsilon$ again. Since ϵ was arbitrary this shows that f is continuous at $x = 0$.

Now consider any $x \neq 0$. Let $\epsilon = |x|/2$, noting that $\epsilon > 0$ since $x \neq 0$. Now consider any $\delta > 0$.

Case: $x \in \mathbb{Q}$. Then $f(x) = 0$ but there is clearly an irrational y close enough to x so that $|y - x| < \min(\epsilon, \delta)$, and hence both $|y - x| < \epsilon$ and $|y - x| < \delta$. We also have that $f(y) = y$. We then have that

$$2\epsilon = |x| \le |y - x| + |y| < \epsilon + |y|$$

so that

$$\epsilon < |y| = |f(y)| = |f(y) - 0| = |f(y) - f(x)| \; .$$

Case: $x \notin \mathbb{Q}$. Then $f(x) = x$, and there is clearly a rational y close enough to x that $|y - x| < \delta$. We then also have $f(y) = 0$ so that

$$|f(y) - f(x)| = |0 - x| = |x| = 2\epsilon > \epsilon$$

since $\epsilon > 0$.

Hence in either case there is a y such that $|y - x| < \delta$ but $|f(y) - f(x)| \ge \epsilon$. This suffices to show that f is not continuous at x. □

Exercise 18.7

(a) Suppose that $f : \mathbb{R} \to \mathbb{R}$ is "continuous from the right," that is

$$\lim_{x \to a^+} f(x) = f(a) \, ,$$

for each $a \in \mathbb{R}$. Show that f is continuous when considered as a function from $\mathbb{R}_l$ to $\mathbb{R}$.

(b) Can you conjecture what functions $f : \mathbb{R} \to \mathbb{R}$ are continuous when considered as maps from $\mathbb{R}$ to $\mathbb{R}_l$? We shall return to this question in Chapter 3.

Solution:

Lemma 18.7.1. *In the topology $\mathbb{R}_l$, every basis element is both open and closed.*

Proof. Consider any basis element $B = [a, b)$, which is clearly open since basis elements are always open. We then have that the complement of this set is $C = \mathbb{R} - B = (-\infty, a) \cup [b, \infty)$. We claim that this complement is also open so that B is closed by definition. To see this, define the sets $C_n = [a - n - 1, a - n + 1) \cup [b + n - 1, b + n + 1)$ for $n \in \mathbb{Z}_+$. Clearly each C_n is open since it is the union of two basis elements. It is also trivial to show that $C = \bigcup_{n \in \mathbb{Z}_+} C_n$, which is then also open since it is a union of open sets. □

Lemma 18.7.2. *The only open sets in the standard topology on $\mathbb{R}$ that are both open and closed are $\varnothing$ and $\mathbb{R}$ itself.*

Proof. First, clearly both $\varnothing$ and $\mathbb{R}$ are both open and closed since they are compliments of each other and are both open by the definition of a topology. Now suppose that U is a nonempty subset of $\mathbb{R}$ that is both open and closed. Suppose also that $U \neq \mathbb{R}$ so that $U \subsetneq \mathbb{R}$ and hence there is a $y \in \mathbb{R}$ where $y \notin U$. We show that the existence of such a U results in a contradiction, which of course shows the desired result since it implies that $U = \mathbb{R}$ if $U \neq \varnothing$. Since U is nonempty we have that there is an $x \in U$ and it must be that $x \neq y$ since $x \in U$ but $y \notin U$.

If $x < y$ then define the set $A = \{z > x \mid z \notin U\}$. Clearly we have that A is nonempty since $y \in A$, and that x is a lower bound of A. It then follows that A has a largest lower bound a since this is a fundamental property of $\mathbb{R}$. It could be that $a \in U$ or $a \notin U$. In the former case we have that any basis element (c, d) containing a is not a subset of U. To see this, we have that $c < a < d$, which means that d is not a lower bound of A since a is the largest lower bound. Hence there is a $z \in A$ where $d > z$. We then have $c < a \leq z < d$ (noting that $a \leq z$ since a is a lower bound of A) so $z \in (c, d)$ and $z \in A$ so that $z \notin U$. Hence (c, d) is not a subset of U, which contradicts the fact that U is open since the basis element (c, d) was arbitrary.

In the latter case where $a \notin U$ then it has to be that $x < a$ since x is a lower bound of A and a is the largest lower bound (and it cannot be that $a = x$ since $x \in U$ but $a \notin U$). We clearly have that $a \in \mathbb{R} - U$, which is open since U is closed. Now consider any basis element (c, d) containing a so that $c < a < d$. Let $b = \max(x, c)$ so that $b < a$ and hence there is a real z where $c \leq b < z < a < d$ and hence $z \in (c, d)$. Now, since $z < a$ it has to be that $z \notin A$ since otherwise a would not be a lower bound of A. We also have that $x \leq b < z$ so that it has to be that $z \in U$ since otherwise it would be that $z \in A$. Thus $z \notin \mathbb{R} - U$, which shows that (c, d) is not a subset of $\mathbb{R} - U$ since $z \in (c, d)$. Since (c, d) was an arbitrary basis element, this contradicts the fact that $\mathbb{R} - U$ is open.

It was thus shown that in either case a contradiction arises. Analogous arguments also show contradictions when $x > y$, this time using the set $A = \{z < x \mid z \notin U\}$ and its least upper bound. Hence it has to be that $U = \mathbb{R}$, which shows the desired result. $\square$

Main Problem.

(a) Recall that by the definition of the one-sided limit, $f : \mathbb{R} \to \mathbb{R}$ is continuous from the right if, for every $a \in \mathbb{R}$ and every $\epsilon > 0$, there is a $\delta > 0$ such that $|f(x) - f(a)| < \epsilon$ for every $x > a$ where $|x - a| < \delta$.

Proof. So suppose that f is continuous from the right and consider any $a \in \mathbb{R}$. Let V be neighborhood of $f(a)$ in $\mathbb{R}$. Then there is a basis element (c, d) of $\mathbb{R}$ that contains $f(a)$ and is a subset of V. Hence $c < f(a) < d$, so let $\epsilon = \min[f(a) - c, d - f(a)]$ so that clearly $\epsilon > 0$ and if $|y - f(a)| < \epsilon$, then $y \in (c, d)$ so that also $y \in V$. Now, since f is continuous from the right, there is a $\delta > 0$ such that $|f(x) - f(a)| < \epsilon$ for every $x > a$ where $|x - a| < \delta$. So let $U = [a, a + \delta)$ which is clearly a basis element of $\mathbb{R}_l$ and contains a so that it is a neighborhood of a.

Now consider any $y \in f(U)$ so that there is an $x \in U$ where $y = f(x)$. If $x = a$ then clearly $|f(x) - f(a)| = |f(a) - f(a)| = |0| = 0 < \epsilon$ so that $f(x) \in V$. If $x \neq a$ then it has to be that $x > a$ and also that $|x - a| = x - a < \delta$ since $U = [a, a + \delta)$. It then follows that $|f(x) - f(a)| < \epsilon$ so that again $f(x) \in V$. Hence in both cases $y = f(x) \in V$, which shows that $f(U) \subset V$ since y was arbitrary. We have thus shown part (4) of Theorem 18.1, from which the topological continuity of f follows. $\square$

(b) We claim that only constant functions are continuous from $\mathbb{R}$ to $\mathbb{R}_l$.

Proof. First, it was shown in Theorem 18.2a that constant functions are always continuous regardless of the topologies. Hence we must show that any continuous function from $\mathbb{R}$ to $\mathbb{R}_l$ is constant. So suppose that f is such a function. Now consider any real x where $x \neq 0$. Clearly if $f(x) = f(0)$ then f is a constant function since x was arbitrary. So suppose that this is not the case so that $f(x) \neq f(0)$. Without loss of generality we can assume that $f(0) < f(x)$. So consider the basis element $B = [f(0), f(x))$ of $\mathbb{R}_l$, which clearly contains $f(0)$ but not $f(x)$.

Since f is continuous and B is both open and closed by Lemma 18.7.1, it follows from the definition of continuity and from Theorem 18.1 part (3) that $f^{-1}(B)$ must be both open and closed in $\mathbb{R}$. However the only sets that are both open and closed in $\mathbb{R}$ are $\varnothing$ and $\mathbb{R}$ itself by Lemma 18.7.2. Thus either $f^{-1}(B) = \varnothing$ or $f^{-1}(B) = \mathbb{R}$. It cannot be that $f^{-1}(B) = \varnothing$ since we have that $f(0) \in B$ so that $0 \in f^{-1}(B)$. Hence it must be that $f^{-1}(B) = \mathbb{R}$, but then we would have $x \in f^{-1}(B)$ so that $f(x) \in B$, which we know it not the case. We therefore have a contradiction so that it must be that $f(x) = f(0)$ so that f is constant. $\square$

Lastly, we claim that the only functions that are continuous from $\mathbb{R}_l$ to $\mathbb{R}_l$ are those that are *continuous and non-decreasing from the right*. For a function $f : \mathbb{R} \to \mathbb{R}$ this means that for every $x \in \mathbb{R}$ and every $\epsilon > 0$ there is a $\delta > 0$ such that $|f(y) - f(x)| < \epsilon$ and $f(y) \geq f(x)$ for every $x \leq y < x + \delta$.

Proof. First we show that such functions are in fact continuous. So suppose that f is continuous and non-decreasing from the right and consider any real x. Let V be any neighborhood of $f(x)$ so that there is a basis element $B = [c, d)$ containing $f(x)$ such that $B \subset V$. Let $\epsilon = d - f(x)$ so that $\epsilon > 0$ since $f(x) < d$. Hence there is a $\delta > 0$ such that $x < y < x + \delta$ implies that $|f(y) - f(x)| < \epsilon$ and $f(y) \geq f(x)$. We then have that $U = [x, x + \delta)$ is a basis element and therefore an open set of $\mathbb{R}_l$ that contains x. Consider any $z \in f(U)$ so that $z = f(y)$ for some $y \in U$. Then $x \leq y < x + \delta$ so that $z = f(y) \geq f(x)$ and $|z - f(x)| = |f(y) - f(x)| < \epsilon$. It then follows that $0 \leq z - f(x) < \epsilon$ so that $c \leq f(x) \leq z < f(x) + \epsilon = d$, and hence $z \in [c, d) = B$. Thus also $z \in V$ since $B \subset V$. This shows that $f(U) \subset V$ since z was arbitrary, and hence that f is continuous by Theorem 18.1.

Now we show that a continuous function *must* be continuous and non-decreasing from the right by showing the contrapositive. So suppose that f is not continuous and non-decreasing from the right. Then there exists a real x and an $\epsilon > 0$ such that, for any $\delta > 0$, there is a $x \leq y < x + \delta$ where $f(y) < f(x)$ or $|f(y) - f(x)| \geq \epsilon$. Clearly we have that $V = [f(x), f(x) + \epsilon)$ is basis element and therefore open set of $\mathbb{R}_l$ that contains $f(x)$. Consider any neighborhood U of x so that there is a basis element $B = [a, b)$ containing x where $B \subset U$. Then $x < b$ so that $\delta = b - x > 0$. It then follows that there is a $x \leq y < x + \delta = b$ such that $f(y) < f(x)$ or $|f(y) - f(x)| \geq \epsilon$. Clearly we have that $y \in B$ so that also $y \in U$ and $f(y) \in f(U)$. However, if $f(y) < f(x)$ then clearly $f(y) \notin V$. On the other hand if $f(y) \geq f(x)$ then it has to be that $|f(y) - f(x)| \geq \epsilon$. Then we have that $f(y) - f(x) \geq 0$ so that $f(y) - f(x) = |f(y) - f(x)| \geq \epsilon$, and hence $f(y) \geq f(x) + \epsilon$ so that again $f(y) \notin V$. This suffices to show that $f(U)$ is not a subset of V, which shows that f is not continuous by Theorem 18.1 since U was an arbitrary neighborhood of x. $\square$

Exercise 18.8

Let Y be an ordered set in the order topology. Let $f, g : X \to Y$ be continuous.

(a) Show that the set $\{x \mid f(x) \leq g(x)\}$ is closed in X.

(b) Let $h : X \to Y$ be the function

$$h(x) = \min\{f(x), g(x)\}.$$

Show that h is continuous. [Hint: Use the pasting lemma.]

Solution:

(a) First let $C = \{x \in X \mid f(x) \leq g(x)\}$ so that we must show that C is closed in X.

Proof. We prove this by showing that the complement $X - C$ is open in X. So first let S be the set of all $y \in Y$ where y has an immediate successor, and denote that successor by $y + 1$. Then clearly $y + 1$ is well defined for all $y \in S$. Now define

$$A_{>y} = \{z \in Y \mid z > y\} \qquad\qquad A_{<y} = \{z \in Y \mid z < y + 1\}$$

for $y \in S$. As these are both rays in the order topology Y, they are both basis elements and therefore open. It then follows that $f^{-1}(A_{>y})$ and $g^{-1}(A_{<y})$ are both open in X since f and g are continuous. Hence their intersection $U_y = f^{-1}(A_{>y}) \cap g^{-1}(A_{<y})$ is also open in X.

Similarly the rays

$$B_{>y} = \{z \in Y \mid z > y\} \qquad\qquad B_{<y} = \{z \in Y \mid z < y\}$$

for $y \in Y$ are also open so that the intersection $V_y = f^{-1}(B_{>y}) \cap g^{-1}(B_{<y})$ is open in X. Then clearly the union of unions

$$D = \bigcup_{y \in S} U_y \cup \bigcup_{y \in Y} V_y$$

is also open in X. We claim that $X - C = D$ so that the complement is open in X and hence C is closed as desired.

($\subset$) First consider any $x \in X - C$ so that clearly $f(x) > g(x)$. If $g(x)$ has an immediate successor $g(x) + 1$ then $g(x) \in S$ and we have $f(x) \in A_{>g(x)}$ so that $x \in f^{-1}(A_{>g(x)})$. We also have that $g(x) \in A_{<g(x)}$ since $g(x) < g(x) + 1$, and hence $x \in g^{-1}(A_{<g(x)})$. It then follows that $x \in U_{g(x)}$ and hence $\bigcup_{y \in S} U_y$ and $x \in D$ since $g(x) \in S$. If $g(x)$ does not have an immediate successor then there must be a $y \in Y$ where $g(x) < y < f(x)$. We then have that clearly $f(x) \in B_{>y}$ and $g(x) \in B_{<y}$ so that $x \in f^{-1}(B_{>y})$ and $x \in g^{-1}(B_{<y})$. Thus $x \in V_y$ so that $x \in \bigcup_{y \in Y} V_y$ and $x \in D$. This shows that $X - C \subset D$ since either way $x \in D$ and x was arbitrary.

($\supset$) Now suppose that $x \in D$. If $x \in \bigcup_{y \in S} U_y$ when there is a $y \in S$ where $x \in U_y$. Hence $x \in f^{-1}(A_{>y})$ and $x \in g^{-1}(A_{<y})$ so that $f(x) \in A_{>y}$ and $g(x) \in A_{<y}$. From this it follows that $f(x) > y$ and $g(x) < y + 1$. Then it has to be that $g(x) \leq y$ so that $f(x) > y \geq g(x)$. If $x \in \bigcup_{y \in Y} V_y$ then there is a $y \in Y$ where $x \in V_y$. Hence $x \in f^{-1}(B_{>y})$ and $x \in g^{-1}(B_{<y})$ so that $f(x) \in B_{>y}$ and $g(x) \in B_{<y}$. It then clearly follows that $f(x) > y$ and $g(x) < y$ so that $f(x) > y > g(x)$. Therefore in either case we have $f(x) > g(x)$ so that $x \in X - C$. This of course shows that $X - C \supset D$ since x was arbitrary. $\square$

(b)

Proof. Let $A = \{x \in X \mid f(x) \leq g(x)\}$ and $B = \{x \in X \mid g(x) \leq f(x)\}$, which are clearly both closed by part (a). It is easy to see that $X = A \cup B$. First, clearly $X \supset A \cup B$ since both $A \subset X$ and $B \subset X$. Then, for any $x \in X$, it has to be that either $f(x) \leq g(x)$ or $f(x) > g(x)$ since $<$ is a total order on Y. In the former case of course $x \in A$, and in the latter $x \in B$ so that either way $x \in A \cup B$. Hence $X \subset A \cup B$. It is also easy to see that $f(x) = g(x)$ for every $x \in A \cap B$. For any such x, we have that $x \in A$ so that $f(x) \leq g(x)$, and $x \in B$ so that $g(x) \leq f(x)$. From this it clearly must be that $f(x) = g(x)$.

Since f and g are continuous, it then follows from the pasting lemma that the function

$$h(x) = \begin{cases} f(x) & x \in A \\ g(x) & x \in B \end{cases}$$

for $x \in X$ is continuous as well. Based on the definitions of A and B it is then easy to see and trivial to show that $h(x) = \min\{f(x), g(x)\}$ for all $x \in X$, which of course shows the desired result. $\square$

Exercise 18.9

Let $\{A_\alpha\}$ be a collection of subsets of X; let $X = \bigcup_\alpha A_\alpha$. Let $f : X \to Y$; suppose that $f \restriction A_\alpha$, is continuous for each α.

(a) Show that if the collection $\{A_\alpha\}$ is finite and each set A_α is closed, then f is continuous.

(b) Find an example where the collection $\{A_\alpha\}$ countable and each A_α is closed, but f is not continuous.

(c) An indexed family of sets $\{A_\alpha\}$ is said to be ***locally finite*** if each point x of X has a neighborhood that intersects A_α for only finitely many values of α. Show that if the family $\{A_\alpha\}$ is locally finite and each A_α is closed, then f is continuous.

Solution:

(a)

Proof. We show using induction that f is continuous for any collection $\{A_\alpha\}_{\alpha=1}^n$, for any $n \in \mathbb{Z}_+$, where each A_α is closed. This of course shows the desired result since the collection is $\{A_\alpha\}_{\alpha=1}^n$ for some $n \in \mathbb{Z}_+$ if it is finite. So first, for $n = 1$, we have that $A_1 = \bigcup_{\alpha=1}^n A_\alpha = X$ so that of course $f = f \restriction X = f \restriction A_1$ is continuous.

Now suppose that f is continuous for any collection of size n and suppose we have the collection $\{A_\alpha\}_{\alpha=1}^{n+1}$ of size $n+1$. Let $A = \bigcup_{\alpha=1}^n A_\alpha$, which is closed by Theorem 17.1 since each A_α is closed and it is a finite union, and let $B = A_{n+1}$ so that B is also closed. We then have that $A \cup B = \bigcup_{\alpha=1}^n A_\alpha \cup A_{n+1} = \bigcup_{\alpha=1}^{n+1} A_\alpha = X$. We know that $g = f \restriction B = f \restriction A_{n+1}$ is continuous. Considering the set A as a subspace of X, then each A_α for $\alpha \in \{1, \ldots, n\}$ is closed in A by Theorem 17.2 since they are subsets of A and closed in X. Since by definition $\bigcup_{\alpha=1}^n A_\alpha = A$, it follows from the induction hypothesis that $f' = f \restriction A$ is continuous. Clearly also for any $x \in A \cap B$ we have that $x \in A$ and $x \in B = A_{n+1}$ so that $f'(x) = (f \restriction A)(x) = f(x) = (f \restriction A_{n+1})(x) = g(x)$.

Then by the pasting lemma the function $h : X \to Y$ defined by

$$h(x) = \begin{cases} f'(x) & x \in A \\ g(x) & x \in B \end{cases}$$

is continuous. However, consider any $x \in X$. If $x \in A$ then $h(x) = f'(x) = (f \restriction A)(x) = f(x)$. Similarly if $x \in B$ then $h(x) = g(x) = (f \restriction B)(x) = f(x)$ as well. This suffices to show that $h = f$ since x was arbitrary. Thus f is continuous, which completes the induction. □

(b) Consider the standard topology on $\mathbb{R}$ and define the countable collection of set $\{A_n\}$ by

$$A_n = \begin{cases} (-\infty, 0] & n = 1 \\ [1, \infty) & n = 2 \\ \left[\frac{1}{n-1}, \frac{1}{n-2}\right] & n > 2 \end{cases}$$

for $n \in \mathbb{Z}_+$. Also define $f : \mathbb{R} \to \mathbb{R}$ by

$$f(x) = \begin{cases} 1 & x \leq 0 \\ 0 & x > 0 \end{cases}$$

for real x. We claim that this collection and function have the desired properties.

Proof. First, it is trivial to show that the collection covers all of $\mathbb{R}$, i.e. that $\bigcup_{n=1}^\infty A_n = \mathbb{R}$. It is also obvious by this point that each A_n is closed in the standard topology. Clearly f is not a continuous function since there is a discontinuity at $x = 0$, which is trivial to prove. Lastly, consider any $n \in \mathbb{Z}_+$. If $n = 1$ then for any $x \in A_n = A_1 = (-\infty, 0]$ we have that $x \leq 0$ and hence $f(x) = 1$. Likewise if $n = 2$ then for any $x \in A_n = A_2 = [1, \infty)$ it follows that $x \geq 1 > 0$, and hence $f(x) = 0$. Lastly, if $n > 2$ then for any $x \in A_n = [1/(n-1), 1/(n-2)]$ we have that $0 < 1/(n-1) \leq x$ so that $f(x) = 0$ again. Thus in all cases $f \restriction A_n$ is constant and therefore continuous. This shows the desired properties. □

(c)

Proof. Consider any $x \in X$ so that there is a neighborhood U' of x that intersects a finite subcollection $\{A_k\}_{k=1}^n$ of the full collection $\{A_\alpha\}$. Consider $A = \bigcup_{k=1}^n A_k$ as a subspace of X, from which it follows from Theorem 17.2 that each A_k is closed in A since it is a subset of A and closed in X. It is then easy to show that $U' \subset A$. It also follows from part (a) that $f \restriction A$ is continuous with the domain being the subspace topology on A.

Now consider any neighborhood V of $f(x)$, noting that of course $x \in A$ since $x \in U'$ and $U' \subset A$. Thus $f(x)$ is in the range of $f \restriction A$ so that there is a neighborhood U_A of x in the subspace topology such that $(f \restriction A)(U_A) \subset V$ by Theorem 18.1 since $f \restriction A$ is continuous. Since U_A is open in the subspace topology, there is an open set U_X in X where $U_A = A \cap U_X$. Now let $U = U' \cap U_X$, which is open in X since U' and U_X are both open in X. Then also $x \in U_X$ since $x \in U_A$ and $U_A = A \cap U_X$, and hence $x \in U$ since also $x \in U'$ and $U = U' \cap U_X$. Thus U is a neighborhood of x in X.

Let z be any element of $f(U)$ so that $z = f(y)$ for some $y \in U$. Then $y \in U'$ and $y \in U_X$ since $U = U' \cap U_X$. Then also $y \in A$ since $U' \subset A$, and hence $y \in A \cap U_X = U_A$. From this it follows that $z = f(y) = (f \restriction A)(y) \in (f \restriction A)(U_A)$ so that $z \in V$ since $(f \restriction A)(U_A) \subset V$. Since z was arbitrary, this shows that $f(U) \subset V$, which in turn shows that f is continuous by Theorem 18.1 since V was an arbitrary neighborhood of $f(x)$ and x was an arbitrary element of X. □

We note that the example in part (b) is not locally finite since any neighborhood of $x = 0$ intersects infinitely many A_n in the collection. This fact is easy to see and would be easy to prove formally, though a bit tedious.

Exercise 18.10

Let $f : A \to B$ and $g : C \to D$ be continuous functions. Let us define a map $f \times g : A \times C \to B \times D$ by the equation

$$(f \times g)(a \times c) = f(a) \times g(c)\,.$$

Show that $f \times g$ is continuous.

Solution:

Proof. Consider any $x \times y \in A \times C$ and any neighborhood V of $(f \times g)(x \times y)$ in $B \times D$. Since V is open in $B \times D$, there is a basis element $U_B \times U_D$ of the product topology that contains $(f \times g)(x \times y)$ where $U_B \times U_D \subset V$. Then U_B and U_D are open in B and D, respectively. Since f is continuous, we then have that $U_A = f^{-1}(U_B)$ is open in A. Likewise $U_C = g^{-1}(U_D)$ is open in C since g is continuous. Then the set $U = U_A \times U_C$ is a basis element of the product topology $A \times C$ and therefore open.

Since $U_B \times U_D$ contains $(f \times g)(x \times y) = f(x) \times g(y)$ we have that $f(x) \in U_B$ and $g(y) \in U_D$. From this it follows that $x \in f^{-1}(U_B) = U_A$ and $y \in g^{-1}(U_D) = U_C$. Therefore $x \times y \in U_A \times U_C = U$ so that U is a neighborhood of $x \times y$ in $A \times C$. Now consider any $w \times z \in (f \times g)(U)$ so that there is an $x' \times y' \in U = U_A \times U_C$ where $w \times z = (f \times g)(x' \times y') = f(x') \times g(y')$. Hence $w = f(x')$ and $x' \in U_A = f^{-1}(U_B)$ so that $w = f(x') \in U_B$. Similarly $z = g(y')$ and $y' \in U_C = g^{-1}(U_D)$ so that $z = g(y') \in U_D$. Thus $w \times z \in U_B \times U_D$ so that also $w \times z \in V$ since $U_B \times U_D \subset V$. This shows that $(f \times g)(U) \subset V$ since $w \times z$ was arbitrary.

This suffices to show that $f \times g$ is continuous by Theorem 18.1 as desired. □

Exercise 18.11

Let $F : X \times Y \to Z$. We say that F is ***continuous in each variable separately*** if for each y_0 in Y, the map $h : X \to Z$ defined by $h(x) = F(x \times y_0)$ is continuous, and for each $x_0 \in X$, the map $k : Y \to Z$ defined by $k(y) = F(x_0 \times y)$ is continuous. Show that if F is continuous, then F is continuous in each variable separately.

Solution:

Proof. To show that F is continuous in x, consider any $y_0 \in Y$ and define $h : X \to Z$ by $h(x) = F(x \times y_0)$. Now consider any $x \in X$ and any neighborhood V of $h(x) = F(x \times y_0)$. Then V is an open set containing $h(x) = F(x \times y_0)$ so that it is a neighborhood of $F(x \times y_0)$. Since F is continuous, this means that there is neighborhood U' of $x \times y_0$ in $X \times Y$ such that $F(U') \subset V$ by Theorem 18.1. It then follows that there is a basis element $U_X \times U_Y$ of $X \times Y$ containing $x \times y_0$ where $U_X \times U_Y \subset U'$. Since $X \times Y$ is a product topology, we have that U_X is open in X and U_Y is open in Y. Then, since $x \times y_0 \in U_X \times U_Y$ we have that $x \in U_X$ and $y_0 \in U_Y$ so that U_X is a neighborhood of x in X.

So consider any $z \in h(U_X)$ so that $z = h(x')$ for some $x' \in U_X$. Then $x' \times y_0 \in U_X \times U_Y$ so that also $x' \times y_0$ in U' since $U_X \times U_Y \subset U'$. It then also follows that $z = h(x') = F(x' \times y_0) \in F(U')$ so that $z \in V$ since $F(U') \subset V$. This shows that $h(U_X) \subset V$ since z was arbitrary. It then follows that h is continuous by Theorem 18.1.

The proof that F is continuous in y is directly analogous. □

Exercise 18.12

Let $F : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ by defined by the equation

$$F(x \times y) = \begin{cases} xy/(x^2 + y^2) & \text{if } x \times y \neq 0 \times 0. \\ 0 & \text{if } x \times y = 0 \times 0. \end{cases}$$

(a) Show that F is continuous in each variable separately.

(b) Compute the function $g : \mathbb{R} \to \mathbb{R}$ defined by $g(x) = F(x \times x)$.

(c) Show that F is not continuous.

Solution:

(a)

Proof. It is easy to see that F is continuous in x. For any real y_0 we generally have that

$$h(x) = F(x \times y_0) = \frac{xy_0}{x^2 + y_0^2}$$

so long as one of x and y_0 are nonzero. If $y_0 = 0$ then $x = 0$ implies that $x \times y_0 = 0 \times 0$ so that $h(x) = F(0 \times 0) = 0$ by definition. If $x \neq 0$ then we have $h(x) = 0/x^2 = 0$ again. Thus h is the constant function $h(x) = 0$ and so is continuous when $y_0 = 0$. If $y_0 \neq 0$ then $y_0^2 > 0$ so that $x^2 + y_0^2 > 0$ since also $x \geq 0$. Thus the denominator is never zero that the $h(x)$ is given by the expression above, which is continuous by elementary calculus. Hence h is always continuous. The same arguments show that F is continuous in y as well. □

(b) We clearly have

$$g(x) = F(x \times x) = \begin{cases} \frac{x^2}{x^2+x^2} = \frac{x^2}{2x^2} = \frac{1}{2} & x \neq 0 \\ 0 & x = 0\,. \end{cases}$$

(c)

Proof. First consider the function $f : \mathbb{R} \to \mathbb{R} \times \mathbb{R}$ defined simply by $f(x) = x \times x$. This function is clearly continuous by Theorem 18.4 since it can be expressed as $f(x) = f_1(x) \times f_2(x)$ where the identical functions $f_1(x) = f_2(x) = x$ are obviously continuous. Then $g = F \circ f$, where g is the function from part (b) since we have $g(x) = F(x \times x) = F(f(x))$ for any real x. Now, clearly g as calculated in part (b) has a discontinuity at $x = 0$ so that it is not continuous. It then follows from Theorem 18.2c that either F or f is not continuous since $g = F \circ f$. As we know that the trivial function f is continuous, it must then be that F is not as desired. □

Exercise 18.13

Let $A \subset X$; let $f : A \to Y$ be continuous; let Y be Hausdorff. Show that if f may be extended to a continuous function $g : \overline{A} \to Y$, then g is uniquely determined by f.

Solution:

Proof. Suppose that g_1 and g_2 are both continuous functions from $\overline{A}$ to Y that extend f so that $g_1(x) = g_2(x) = f(x)$ for all $x \in A$. Clearly $g_1 = g_2$ if and only if $g_1(x) = g_2(x)$ for all $x \in \overline{A}$. So suppose that this is *not* the case so that there is an $x_0 \in \overline{A}$ where $g_1(x_0) \neq g_2(x_0)$. Since Y is a Hausdorff space and $g_1(x_0)$ and $g_2(x_0)$ are distinct, there are disjoint neighborhoods V_1 and V_2 of $g_1(x_0)$ and $g_2(x_0)$, respectively. Then there are also neighborhoods U_1 and U_2 of x_0 such that $g_1(U_1) \subset V_1$ and $g_2(U_2) \subset V_2$ by Theorem 18.1 since both g_1 and g_2 are continuous.

Now let $U = U_1 \cap U_2$ so that U is also a neighborhood of x_0. Since $x_0 \in \overline{A}$, it follows that U intersects A so that there is a $y \in U$ where also $y \in A$ by Theorem 17.5. Since $y \in A$ we have that $g_1(y) = g_2(y) = f(y)$. We also have that $y \in U_1$ and $y \in U_2$ since $U = U_1 \cap U_2$. Thus $g_1(y) \in g_1(U_1)$ so that $f(y) = g_1(y) \in V_1$ since $g_1(U_1) \subset V_1$. Similarly $f(y) = g_2(y) \in V_2$, but then we have that $f(y) \in V_1 \cap V_2$, which contradicts the fact that V_1 and V_2 are disjoint! Hence it must be that $g_1 = g_2$, which shows uniqueness. □

§19 The Product Topology

Exercise 19.1

Prove Theorem 19.2

Solution:

Let $\mathcal{C}$ be the collection of sets that are alleged to be a basis for the box or product topologies in Theorem 19.2.

Proof. We show that $\mathcal{C}$ is a basis of the box or product topology using Lemma 13.2. First, it is easy to see that $\mathcal{C}$ is a collection of open sets. Consider any $B \in \mathcal{C}$ so that $B = \prod B_\alpha$ where each $B_\alpha \in \mathcal{B}_\alpha$ (for a finitely many $\alpha \in J$ and $B_\alpha = X_\alpha$ for the rest in the product topology). Since each B_α is a basis element of X_α (or X_α itself), they are open so that B is a basis element of the box or product topology by definition and therefore open. Note that the basis for the product topology is given directly by Theorem 19.1.

Now suppose that U is an any open set of the box topology and consider any $x \in U$. Then it follows that there is a basis element $\prod_{\alpha \in J} U_\alpha$ of the box or product topology containing x where $\prod_{\alpha \in J} U_\alpha \subset U$. Thus each U_α is an open set of X_α (or $U_\alpha = X_\alpha$ for all but finitely many $\alpha \in J$ for the product topology). Also $x \in \prod_{\alpha \in J} U_\alpha$ so that $x = (x_\alpha)_{\alpha \in J}$ where each $x_\alpha \in U_\alpha$. It then follows that there is basis element $B_\alpha \in \mathcal{B}_\alpha$ of X_α containing x_α where $B_\alpha \subset U_\alpha$ (for $U_\alpha = X_\alpha$ we simply set $B_\alpha = X_\alpha$ as well).

Then clearly $x \in \prod_{\alpha \in J} B_\alpha$ and $\prod_{\alpha \in J} B_\alpha \in \mathcal{C}$. Consider also any $y \in \prod_{\alpha \in J} B_\alpha$ so that $y = (y_\alpha)_{\alpha \in J}$ where each $y_\alpha \in B_\alpha$. Then also each $y_\alpha \in U_\alpha$ since $B_\alpha \subset U_\alpha$. This suffices to show that $y \in \prod_{\alpha \in J} U_\alpha \subset U$. Since y was arbitrary this shows that $\prod_{\alpha \in J} B_\alpha \subset U$. Therefore $\mathcal{C}$ is a basis of the box topology by Lemma 13.2. □

Exercise 19.2

Prove Theorem 19.3.

Solution:

Proof. The basis of the box or product topologies on $\prod A_\alpha$ is the collection of sets $\prod V_\alpha$, where each V_α is open in A_α and, in the case of the product topology, $V_\alpha = A_\alpha$ for all but finitely many $\alpha \in J$ (by Theorem 19.1). Denote this basis collection by $\mathcal{C}$. By Lemma 16.1, the collection

$$\mathcal{B}_A = \left\{ B \cap \prod A_\alpha \mid B \in \mathcal{B} \right\}$$

is a basis of the subspace topology on $\prod A_\alpha$, where $\mathcal{B}$ is the basis of $\prod X_\alpha$. To prove that $\prod A_\alpha$ is a subspace of $\prod X_\alpha$, it therefore suffices to show that $\mathcal{C} = \mathcal{B}_A$.

($\subset$) First consider any element $B \in \mathcal{C}$ so that $B = \prod V_\alpha$ for open sets V_α in A_α (and $V_\alpha = A_\alpha$ for all but finite many $\alpha \in J$ for the product topology). For each $\alpha \in J$, we then have that $V_\alpha = U_\alpha \cap A_\alpha$ for some open set U_α in X_α since A_α is a subspace of X_α. Note that this is true even for those α where $V_\alpha = A_\alpha$ in the product topology since then $V_\alpha = A_\alpha = X_\alpha \cap A_\alpha$. In fact, for these α we need to choose $U_\alpha = X_\alpha$ as will become apparent. We then have the following:

$$\begin{aligned}
x \in B &\Leftrightarrow x \in \prod V_\alpha \\
&\Leftrightarrow \forall \alpha \in J (x_\alpha \in V_\alpha) \\
&\Leftrightarrow \forall \alpha \in J (x_\alpha \in U_\alpha \cap A_\alpha) \\
&\Leftrightarrow \forall \alpha \in J (x_\alpha \in U_\alpha \wedge x_\alpha \in A_\alpha) \\
&\Leftrightarrow \forall \alpha \in J (x_\alpha \in U_\alpha) \wedge \forall \alpha \in J (x_\alpha \in A_\alpha) \\
&\Leftrightarrow x \in \prod U_\alpha \wedge x \in \prod A_\alpha \\
&\Leftrightarrow x \in \left(\prod U_\alpha \right) \cap \left(\prod A_\alpha \right) ,
\end{aligned}$$

Since $U_\alpha = X_\alpha$ for all but a finitely many $\alpha \in J$ for the product topology, we have that $\prod U_\alpha$ is a basis element of $\prod X_\alpha$, i.e. $\prod U_\alpha \in \mathcal{B}$. This shows that $B \in \mathcal{B}_A$ so that $\mathcal{C} \subset \mathcal{B}_A$ since B was arbitrary.

($\supset$) Now suppose that $B \in \mathcal{B}_A$ so that $B = B_X \cap \prod A_\alpha$ for some basis element $B_X \in \mathcal{B}$ of $\prod X_\alpha$. We then have that $B_X = \prod U_\alpha$ where each U_α is an open set of X_α (and $U_\alpha = X_\alpha$ for all but finitely many $\alpha \in J$ for the product topology). Then let $V_\alpha = U_\alpha \cap A_\alpha$ for each $\alpha \in J$, noting that $V_\alpha = X_\alpha \cap A_\alpha = A_\alpha$ when $U_\alpha = X_\alpha$. Following the above chain of logical equivalences in reverse order then shows that $B = \prod V_\alpha$ so that $B \in \mathcal{C}$ since clearly each V_α is open in the subspace topology A_α. Hence $\mathcal{C} \supset \mathcal{B}_A$ since B was arbitrary. □

Exercise 19.3

Prove Theorem 19.4.

Solution:

Proof. Suppose that x and y are distinct points of $\prod X_\alpha$. Then $x = (x_a)$ and $y = (y_\alpha)$ where each $x_\alpha, y_\alpha \in X_\alpha$, and there must be a β where $x_\beta \neq y_\beta$ since $x \neq y$. Thus x_β and y_β are distinct points of X_β, so that there are neighborhoods W_x and W_y of x_β and y_β, respectively, that are disjoint since X_β is a Hausdorff space. So define the sets

$$U_\alpha = \begin{cases} W_x & \alpha = \beta \\ X_\alpha & \alpha \neq \beta \end{cases} \qquad\qquad V_\alpha = \begin{cases} W_y & \alpha = \beta \\ X_\alpha & \alpha \neq \beta \end{cases}$$

so that clearly $x \in \prod U_\alpha$ and $y \in \prod V_\alpha$. Then since each U_α and V_α are open, we have that $\prod U_\alpha$ and $\prod V_\alpha$ are both basis elements of $\prod X_\alpha$ and therefore open. Note that this is true for both the box and product topologies since, in the case of the latter, U_α and V_α are not all of X_α for only one α, namely $\alpha = \beta$. Thus $\prod U_\alpha$ is a neighborhood of x and $\prod V_\alpha$ is a neighborhood of y in $\prod X_\alpha$.

We also assert that $\prod U_\alpha$ and $\prod V_\alpha$ are disjoint, which of course completes the proof that $\prod X_\alpha$ is Hausdorff. To see this, suppose to the contrary that there is a z in both $\prod U_\alpha$ and $\prod V_\alpha$. Then $z = (z_\alpha)$ and in particular we would have that $z_\beta \in U_\beta = W_x$ and $z_\beta \in V_\beta = W_y$. But then $z_\beta \in W_x \cap W_y$, which contradicts the fact that W_x and W_y are disjoint! So it must be that in fact $\prod U_\alpha$ and $\prod V_\alpha$ are disjoint. $\square$

Exercise 19.4

Show that $(X_1 \times \cdots \times X_{n-1}) \times X_n$ is homeomorphic to $X_1 \times \cdots \times X_n$.

Solution:

Proof. First we note that since we are dealing with finite products, the box and product topologies are the same; we shall find it most convenient to use the box topology definition. Also, as there are no intervals involved here, we use the traditional tuple notation using parentheses. So define $f : X_1 \times \cdots \times X_n \to (X_1 \times \cdots \times X_{n-1}) \times X_n$ by

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

It is obvious that this is a bijection, and it is trivial to prove. Also obvious and trivial to prove based on the definition of f is that $f(A_1 \times \cdots \times A_n) = (A_1 \times \cdots \times A_{n-1}) \times A_n$ when each $A_k \subset X_k$.

First we show that f is continuous by showing that the inverse image of every basis element in $(X_1 \times \cdots \times X_{n-1}) \times X_n$ is open in $X_1 \times \cdots \times X_n$. So consider any basis element C of $(X_1 \times \cdots \times X_{n-1}) \times X_n$ and let $U = f^{-1}(C)$ so that of course $f(U) = C$ and $U \subset X_1 \times \cdots \times X_n$. We then have that $C = V' \times V_n$ where V' is open in $X_1 \times \cdots \times X_{n-1}$ and V_n is open in X_n by the definition of the box/product topology. Now consider any $x \in U$ so that $x = (x_k)_{k=1}^n$ and we have that $f(x) = ((x_1, \ldots, x_{n-1}), x_n) \in f(U) = C$. Hence $x' = (x_1, \ldots, x_{n-1}) \in V'$ and $x_n \in V_n$. Since V' is open in $X_1 \times \cdots \times X_{n-1}$ there is a basis element C' containing x' that is a subset of V'. By the definition of the box topology, we then have that $C' = V_1 \times \cdots \times V_{n-1}$ where each V_k is open in X_k.

We then have that $B = V_1 \times \cdots \times V_n$ is a basis element of $X_1 \times \cdots \times X_n$ and also clearly B contains x since $(x_1, \ldots, x_{n-1}) = x' \in C' = V_1 \times \cdots \times V_{n-1}$ and $x_n \in V_n$. Now suppose that $y = (y_k)_{k=1}^n \in B$ so that each $y_k \in V_k$. Then we have that $y' = (y_1, \ldots, y_{n-1}) \in C'$ so that also $y' \in V'$ since $C' \subset V'$. Since also of course $y_n \in V_n$, we have that $(y', y_n) \in V' \times V_n = C$. Also clearly $f(y) = (y', y_n) \in C = f(U)$ so that $y \in U$. Since y was arbitrary this shows that $B \subset U$, which suffices to show that U is open since x was arbitrary. This completes the proof that f is continuous.

Next we show that f^{-1} is continuous, which is a little simpler. Let B be any basis element of $X_1 \times \cdots \times X_n$ so that $B = U_1 \times \cdots \times U_n$ where each U_k is open in X_k by the definition of the box topology. Then we have that $f(B) = (U_1 \times \cdots \times U_{n-1}) \times U_n$. By the definition of the box topology, we then have that $U' = U_1 \times \cdots \times U_{n-1}$ is a basis element of $X_1 \times \cdots \times X_{n-1}$ and is therefore open. Since U_n is also open, we have that $f(B) = U' \times U_n$ is a basis element of $(X_1 \times \cdots \times X_{n-1}) \times X_n$ by the definition of the box/product topology, and is therefore open. Since $f(B) = (f^{-1})^{-1}(B)$ is the inverse image of B under f^{-1}, this shows that f^{-1} is also continuous.

We have shown that both f and f^{-1} are continuous, which proves that f is a homeomorphism by definition. □

Exercise 19.5

One of the implications stated in Theorem 19.6 holds for the box topology. Which one?

Solution:

Example 19.2 gives a function f that is not continuous in the box topology even though all of its constituent functions f_α are continuous. Hence the only implication that can be generally true in the box topology is that f being continuous implies that each f_α is continuous. A proof of this is straightforward.

Proof. As in Theorem 19.6 suppose that $f : A \to \prod_{\alpha \in J} X_\alpha$ be given by

$$f(a) = (f_\alpha(a))_{\alpha \in J} ,$$

where $f_\alpha : A \to X_\alpha$ for each $\alpha \in J$. Here $\prod X_\alpha$ has the box topology. Suppose that f is continuous and consider any $\beta \in J$. We show that f_β is continuous, which of course shows the desired result.

So let V be any open set of X_β and define

$$B_\alpha = \begin{cases} V & \alpha = \beta \\ X_\alpha & \alpha \neq \beta \,. \end{cases}$$

Then, since each B_α is clearly open in X_α, we have that $B = \prod B_\alpha$ is a basis element of the box topology by definition and is therefore open. Hence $U = f^{-1}(B)$ is open in A since f is continuous. We claim that $U = f_\beta^{-1}(V)$, which shows that f_β is continuous since U is open in A and V was an arbitrary open set of X_β.

($\subset$) If $x \in U = f^{-1}(B)$ then of course $f(x) \in B$ so that each $f_\alpha(x) \in B_\alpha$ since $f(x) = (f_\alpha(x))_{\alpha \in J}$ and $B = \prod B_\alpha$. In particular $f_\beta(x) \in B_\beta = V$ so that $x \in f_\beta^{-1}(V)$. Hence $U \subset f_\beta^{-1}(V)$ since x was arbitrary.

($\supset$) If $x \in f_\beta^{-1}(V)$ then $f_\beta(x) \in V = B_\beta$. Since of course every other $f_\alpha(x) \in X_\alpha = B_\alpha$ we have that $f(x) \in \prod B_\alpha = B$. Hence $x \in f^{-1}(B) = U$ so that $f_\beta^{-1}(V) \subset U$ since x was arbitrary. □

Exercise 19.6

Let $\mathbf{x}_1, \mathbf{x}_2, \ldots$ be a sequence of the points of the product space $\prod X_\alpha$. Show that the sequence converges to the point $\mathbf{x}$ if and only if the sequence $\pi_\alpha(\mathbf{x}_1), \pi_\alpha(\mathbf{x}_2), \ldots$ converges to $\pi_\alpha(\mathbf{x})$ for each α. Is this fact true if one uses the box topology instead of the product topology?

Solution:

Proof. ($\Rightarrow$) First suppose that the sequence $\mathbf{x}_1, \mathbf{x}_2, \ldots$ converges to $\mathbf{x}$ and consider any β. Also suppose that U is any neighborhood of $\pi_\beta(\mathbf{x})$. Define

$$B_a = \begin{cases} U & \alpha = \beta \\ X_\alpha & \alpha \neq \beta \end{cases}$$

so that $B = \prod B_\alpha$ is a basis element of $\prod X_\alpha$ since each B_α is open. Note that B is a basis element of both the box and product topologies since possibly $B_\alpha \neq X_\alpha$ for only one α (i.e. for $\alpha = \beta$). We also clearly have that $\mathbf{x} \in B$ so that B is a neighborhood of $\mathbf{x}$ in $\prod X_\alpha$. Since the sequence $\mathbf{x}_1, \mathbf{x}_2, \ldots$ converges to $\mathbf{x}$, we have that there is an $N \in \mathbb{Z}_+$ where $\mathbf{x}_n \in B$ for all $n \geq N$. So consider any such $n \geq N$ so that $\mathbf{x}_n \in B = \prod B_\alpha$. Hence $\pi_\alpha(\mathbf{x}_n) \in B_\alpha$ for all α, and in particular $\pi_\beta(\mathbf{x}_n) \in B_\beta = U$. This suffices to show that the sequence $\pi_\beta(\mathbf{x}_1), \pi_\beta(\mathbf{x}_2), \ldots$ converges to $\pi_\beta(\mathbf{x})$ as desired since U was an arbitrary neighborhood.

($\Leftarrow$) Now suppose that the sequence $\pi_\alpha(\mathbf{x}_1), \pi_\alpha(\mathbf{x}_2), \ldots$ converges to $\pi_\alpha(\mathbf{x})$ for every α. Let U be any neighborhood of $\mathbf{x}$ in $\prod X_\alpha$. Then there is a basis element $B = \prod U_\alpha$ of $\prod X_\alpha$ where $\mathbf{x} \in B$ and $B \subset U$. Since $\prod X_\alpha$ is the product topology, each U_α is open but only a finite number of them are different from X_α. Suppose then that J is the index set of α and that $I \subset J$ is the finite subset where $U_\alpha = X_\alpha$ for all $\alpha \notin I$.

Then for any $\beta \in I$ we have that $\pi_\beta(\mathbf{x}) \in U_\beta$ since $\mathbf{x} \in B = \prod U_\alpha$, hence U_β is a neighborhood of $\pi_\beta(\mathbf{x})$. Then, since $\pi_\beta(\mathbf{x}_1), \pi_\beta(\mathbf{x}_2), \ldots$ converges to $\pi_\beta(\mathbf{x})$, there is an $N_\beta \in \mathbb{Z}_+$ where $\pi_\beta(\mathbf{x}_n) \in U_\beta$ for all $n \geq N_\beta$. So let $N = \max_{\alpha \in I} N_\alpha$, noting that this exists since I is finite. Consider any $n \geq N$ and any $\alpha \in J$. If $\alpha \in I$ then we have that $n \geq N \geq N_\alpha$ so that $\pi_\alpha(\mathbf{x}_n) \in U_\alpha$. If $\alpha \notin J$ then of course we have that $\pi_a(\mathbf{x}_n) \in X_\alpha = U_\alpha$. Hence either way we have that $\pi_\alpha(\mathbf{x}_n) \in U_\alpha$ so that $\mathbf{x}_n \in \prod U_\alpha = B$ and hence also $\mathbf{x}_n \in U$ since $B \subset U$. Since $n \geq N$ was arbitrary and U was an arbitrary neighborhood of $\mathbf{x}$, this shows that $\mathbf{x}_1, \mathbf{x}_2, \ldots$ converges to $\mathbf{x}$ as desired. □

As noted there, the forward direction of the preceding proof works for the product or the box topology. However, then reverse direction was proved only for the product topology, with the critical point being where we took $\max_{\alpha \in I} N_\alpha$, which was only guaranteed to exist since I is finite in the product topology. The provides a hint as to how to construct a counterexample that proves that this direction is not generally true for the box topology.

Proof. Define

$$x_{ij} = \begin{cases} 1 & j \leq i \\ \frac{1}{j-i} & j > i \end{cases}$$

for $i, j \in \mathbb{Z}_+$. Now define a sequence $\mathbf{x}_1, \mathbf{x}_2, \ldots$ in $\prod_{i \in \mathbb{Z}_+} \mathbb{R} = \mathbb{R}^\omega$ by $\pi_i(\mathbf{x}_j) = x_{ij}$. With the box topology on $\mathbb{R}^\omega$ we claim that each coordinate sequence $\pi_i(\mathbf{x}_1), \pi_i(\mathbf{x}_2), \ldots$ converges to 0 but that the sequence $\mathbf{x}_1, \mathbf{x}_2, \ldots$ does not converge to the point $\mathbf{0} = (0, 0, \ldots)$.

First, it is easy to see that each coordinate sequence $\pi_i(\mathbf{x}_1), \pi_i(\mathbf{x}_2), \ldots$ converges to 0 since, for fixed i, there is always an $N \in \mathbb{Z}_+$ large enough such that $j > i$ and $\pi_i(\mathbf{x}_j) = x_{ij} = 1/(j-i)$ is small enough to be within any fixed neighborhood of 0 for all $j \geq N$. To show that the sequence $\mathbf{x}_1, \mathbf{x}_2, \ldots$ does not converge to $\mathbf{0}$ though, consider the neighborhood $U = \prod U_k$ of $\mathbf{0}$ where every $U_k = (-1, 1)$. We note that clearly U is open in the box topology since each U_k is a basis element of $\mathbb{R}$ and therefore open. For any $N \in \mathbb{Z}_+$ we then have that $\pi_N(\mathbf{x}_N) = x_{NN} = 1$ so that clearly $\pi_N(\mathbf{x}_N) \notin (-1, 1) = U_N$ and hence $\mathbf{x}_N \notin \prod U_k = U$. This suffices to show that the sequence does not converge, but it does not even come close to converging since there are actually no points in the sequence that are even in this quite large neighborhood of $\mathbf{0}$! □

Exercise 19.7

Let $\mathbb{R}^\infty$ be the subset of $\mathbb{R}^\omega$ consisting of all sequences that are "eventually zero," that is all sequences $(x_1, x_2, \ldots)$ such that $x_i \neq 0$ for only finitely many values of i. What is the closure of $\mathbb{R}^\infty$ in $\mathbb{R}^\omega$ in the box and product topologies? Justify your answer.

Solution:

First we claim that $\mathbb{R}^\infty$ is dense in $\mathbb{R}^\omega$ in the product topology in the sense that its closure is all of $\mathbb{R}^\omega$.

Proof. We show that any point of $\mathbb{R}^\omega$ is in $\overline{\mathbb{R}^\omega}$. So consider any point $x = (x_n)_{n\in\mathbb{Z}_+} \in \mathbb{R}^\omega$ and any neighborhood U of x. Then there is a basis element $B = \prod U_n$ containing x where $B \subset U$. By the definition of the product topology each U_n is open and $U_n = \mathbb{R}$ for all but finitely many values of n. So let I be a finite subset of $\mathbb{Z}_+$ such that $U_n = \mathbb{R}$ for all $n \notin I$ and U_n is merely just open for $n \in I$.

Consider now the sequence $y = (y_n)_{n\in\mathbb{Z}_+}$ defined by

$$y_n = \begin{cases} x_n & n \in I \\ 0 & n \notin I \end{cases}$$

for $n \in \mathbb{Z}_+$. Since I is finite clearly $y \in \mathbb{R}^\infty$. Also $y_n = x_n \in U_n$ when $n \in I$ since $B = \prod U_n$ contains x. We also have $y_n = 0 \in \mathbb{R} = U_n$ when $n \notin I$ so that either way $y_n \in U_n$ and hence $y \in \prod U_n = B$. Thus also $y \in U$ since $B \subset U$. Since U was an arbitrary neighborhood and U intersects $\mathbb{R}^\infty$ (with y being a point in the intersection), this shows that $x \in \overline{\mathbb{R}^\infty}$ by Theorem 17.5. This of course shows the desired result since x was any element of $\mathbb{R}^\omega$. □

For the box topology, we claim that $\mathbb{R}^\infty$ is already closed.

Proof. We show this by showing that any point not in $\mathbb{R}^\infty$ is not a limit point of $\mathbb{R}^\infty$ so that $\mathbb{R}^\infty$ must already contain all its limit points. So consider any $x = (x_n)_{n\in\mathbb{Z}_+} \notin \mathbb{R}^\infty$ so that $x_n \neq 0$ for infinitely many values of n. Now define the sets

$$U_n = \begin{cases} (-1, 1) & x_n = 0 \\ (x_n/2, 2x_n) & x_n > 0 \\ (2x_n, x_n/2) & x_n < 0 \end{cases}$$

for $n \in \mathbb{Z}_+$. Clearly each U_n is a basis element of $\mathbb{R}$ and is therefore open. Also clearly each $x_n \in U_n$. It therefore follows that $B = \prod U_n$ is a basis element of $\mathbb{R}^\omega$ and is therefore open, and that $x \in B$. Hence B is a neighborhood of x.

Then, for any $y = (y_n)_{n\in\mathbb{Z}_+} \in B$ we have that each $y_n \in U_n$. For infinitely many $n \in \mathbb{Z}_+$ we then have that $x_n \neq 0$ and hence $x_n > 0$ or $x_n < 0$. In the former case $y_n \in U_n = (x_n/2, 2x_n)$ so that $0 < x_n/2 < y_n$. In the latter case $y_n \in U_n = (2x_n, x_n/2)$ so that $y_n < x_n/2 < 0$. Hence either way $y_n \neq 0$ so that $y \notin \mathbb{R}^\infty$ since this is true for infinitely many n. Since $y \in B$ was arbitrary, this shows that B cannot not intersect $\mathbb{R}^\infty$. Therefore x is not a limit point of $\mathbb{R}^\infty$ since B is a neighborhood of x. □

Exercise 19.8

Given sequences $(a_1, a_2, \ldots)$ and $(b_1, b_2, \ldots)$ of reals numbers with $a_i > 0$ for all i, define $h : \mathbb{R}^\omega \to \mathbb{R}^\omega$ by the equation

$$h((x_1, x_2, \ldots)) = (a_1x_1 + b_1, a_2x_2 + b_2, \ldots)\,.$$

Show that if $\mathbb{R}^\omega$ is given the product topology, h is a homeomorphism of $\mathbb{R}^\omega$ with itself. What happens if $\mathbb{R}^\omega$ is given the box topology?

Solution:

Lemma 19.8.1. *Consider the spaces $\prod X_\alpha$ and $\prod Y_\alpha$ in the box topologies over the index set J. If $f : \prod X_\alpha \to \prod Y_\alpha$ is defined by*

$$f((x_\alpha)_{\alpha\in J}) = (f_\alpha(x_\alpha))_{\alpha\in J}$$

and each $f_\alpha : X_\alpha \to Y_\alpha$ is continuous, then f is continuous.

Proof. Consider any basis element $B = \prod V_\alpha$ in $\prod Y_\alpha$ so that each V_α is open in Y_α since we are in the box topology. For each $\alpha \in J$ then define $U_\alpha = f_\alpha^{-1}(V_\alpha)$, which is open in X_α since f_α is continuous. Hence the set $U = \prod U_\alpha$ is a basis element of $\prod X_\alpha$ in the box topology and is therefore open. We claim that $U = f^{-1}(B)$, which shows that f is continuous since U is open and B was arbitrary.

($\subset$) Consider any $\mathbf{x} \in U = \prod U_\alpha$. Then, for any $\alpha \in J$, we have $x_\alpha \in U_\alpha = f_\alpha^{-1}(V_\alpha)$ so that $f(x_\alpha) \in V_\alpha$. Hence $f(\mathbf{x}) = (f_\alpha(x_\alpha))_{\alpha\in J} \in \prod V_\alpha = B$ so that $\mathbf{x} \in f^{-1}(B)$. this shows that $U \subset f^{-1}(B)$ since $\mathbf{x}$ was arbitrary.

($\supset$) Now consider any $\mathbf{x} \in f^{-1}(B)$ so that $f(\mathbf{x}) \in B = \prod V_\alpha$ and hence each $f_\alpha(x_\alpha) \in V_\alpha$ by the definition of f. Then $x_\alpha \in f_\alpha^{-1}(V_\alpha) = U_\alpha$ so that clearly $\mathbf{x} \in \prod U_\alpha = U$. Since $\mathbf{x}$ was arbitrary this shows that $f^{-1}(B) \subset U$ as well. $\square$

Main Problem.

Proof. First note that clearly $h(\mathbf{x}) = (h_1(\mathbf{x}), h_2(\mathbf{x}), \ldots)$ for $\mathbf{x} \in \mathbb{R}^\omega$, where each $h_i : \mathbb{R}^\omega \to \mathbb{R}$ is defined by

$$h_i(\mathbf{x}) = a_i \pi_i(\mathbf{x}) + b_i \,.$$

This can further be broken down as $h_i(\mathbf{x}) = f_i(\pi_i(\mathbf{x})) = (f_i \circ \pi_i)(\mathbf{x})$, where each $f_i : \mathbb{R} \to \mathbb{R}$ is defined by $f_i(x) = a_i x + b_i$. As discussed in the proof of Theorem 19.6, each π_i is continuous and we have that each f_i is continuous by elementary calculus, noting that this is true whether each $a_i > 0$ or not. It then follows from Theorem 18.2c that each $f_i \circ \pi_i = h_i$ is continuous. Then we have that h is continuous by Theorem 19.6 since each coordinate function is continuous and we are using the product topology.

Now define the functions $g_i : \mathbb{R} \to \mathbb{R}$ by $g_i(x) = (x - b_i)/a_i$ for $i \in \mathbb{Z}_+$, noting that this is defined since each $a_i > 0$. Define also the functions $k_i : \mathbb{R}^\omega \to \mathbb{R}$ by $k_i = g_i \circ \pi_i$, and finally define $k : \mathbb{R}^\omega \to \mathbb{R}^\omega$ by $k(\mathbf{x}) = (k_1(\mathbf{x}), k_2(\mathbf{x}), \ldots)$. Now again we have that each π_i and g_i are continuous by the proof of Theorem 19.6 and elementary calculus. Hence $k_i = g_i \circ \pi_i$ and k are continuous by Theorem 18.2c, and Theorem 19.6, respectively, as before.

Now consider any $\mathbf{x} = (x_i)_{i\in\mathbb{Z}_+} \in \mathbb{R}^\omega$ so that we have, for any $i \in \mathbb{Z}_+$,

$$\begin{aligned} k_i(h(\mathbf{x})) &= [g_i \circ \pi_i](h(\mathbf{x})) = g_i(\pi_i(h(\mathbf{x}))) = g_i(h_i(\mathbf{x})) \\ &= g_i([f_i \circ \pi_i](\mathbf{x})) = g_i(f_i(\pi_i(\mathbf{x}))) = g_i(f_i(x_i)) \\ &= \frac{f_i(x_i) - b_i}{a_i} = \frac{(a_i x_i + b_i) - b_i}{a_i} = \frac{a_i x_i}{a_i} \\ &= x_i \,. \end{aligned}$$

Therefore

$$k(h(\mathbf{x})) = (k_1(h(\mathbf{x})), k_2(h(\mathbf{x})), \ldots) = (x_1, x_2, \ldots) = \mathbf{x} \,.$$

We also have that

$$\begin{aligned} h_i(k(\mathbf{x})) &= [f_i \circ \pi_i](k(\mathbf{x})) = f_i(\pi_i(k(\mathbf{x}))) = f_i(k_i(\mathbf{x})) \\ &= f_i([g_i \circ \pi_i](\mathbf{x})) = f_i(g_i(\pi_i(\mathbf{x}))) = f_i(g_i(x_i)) \\ &= a_i g_i(x_i) + b_i = a_i \left(\frac{x_i - b_i}{a_i} \right) + b_i = (x_i - b_i) + b_i \\ &= x_i \,. \end{aligned}$$

for each $i \in \mathbb{Z}_+$ so that

$$h(k(\mathbf{x})) = (h_1(k(\mathbf{x})), h_2(k(\mathbf{x})), \ldots) = (x_1, x_2, \ldots) = \mathbf{x} \,.$$

Since $\mathbf{x}$ was arbitrary, it thus follows from Lemma 2.1 that h is bijective and $k = h^{-1}$. Since we have already shown that h and $k = h^{-1}$ are continuous, this suffices to prove that h is a homeomorphism as desired. □

We claim that h is also a homeomorphism in the box topology.

Proof. First, h is still a bijection as the proof of this above does not depend on the topology at all. However, Theorem 19.6 was used in the proofs that h and h^{-1} are continuous, and we know that this theorem is not generally true for the box topology. On the other hand h can be formulated as $h(\mathbf{x}) = (f_1(x_1), f_2(x_2), \ldots)$, where as before each $f_i(x) = a_i x + b_i$. Since each f_i is continuous by elementary calculus, it follows from Lemma 19.8.1 that h is continuous in the box topology. The same argument applies to the inverse function h^{-1} since $h^{-1}(\mathbf{x}) = (g_1(x_1), g_2(x_2), \ldots)$ and each g_i is continuous. □

Exercise 19.9

Show that the choice axiom is equivalent to the statement that for any indexed family $\{A_\alpha\}_{\alpha \in J}$ of nonempty sets, with $J \neq 0$, the cartesian product

$$\prod_{\alpha \in J} A_\alpha$$

is not empty.

Solution:

Proof. For the following denote the collection $\{A_\alpha\}_{\alpha \in J}$ by $\mathcal{A}$.

($\Rightarrow$) First suppose that the choice axiom is true. Then by Lemma 9.2 there exists a choice function

$$c : \mathcal{A} \to \bigcup_{A \in \mathcal{A}} A$$

where $c(A) \in A$ for each $A \in \mathcal{A}$, noting that this is true since $\mathcal{A}$ is a collection of nonempty sets. Then consider, set $x_\alpha = c(A_\alpha)$ for each $\alpha \in J$ so that $x_\alpha = c(A_\alpha) \in A_\alpha$. Therefore clearly $\mathbf{x} = (x_\alpha)_{\alpha \in J} \in \prod A_\alpha$ so that $\prod A_\alpha$ is not empty.

($\Leftarrow$) Now suppose that $\prod_{\alpha \in J} A_\alpha$ is nonempty for any indexed family $\{A_\alpha\}_{\alpha \in J}$ of nonempty sets when $J \neq \varnothing$. Let $\mathcal{A}$ be a collection of disjoint nonempty sets where $\mathcal{A} \neq \varnothing$. Then the $\{A\}_{A \in \mathcal{A}}$ is a nonempty family of nonempty sets. Hence $\prod_{A \in \mathcal{A}} A$ is nonempty so that there is an $\mathbf{x} = (x_A)_{A \in \mathcal{A}} \in \prod_{A \in \mathcal{A}} A$, and thus $x_A \in A$ for every $A \in \mathcal{A}$. Now let $C = \{x_A\}_{A \in \mathcal{A}}$ so that clearly $C \subset \bigcup \mathcal{A}$.

Consider any $A \in \mathcal{A}$ so that $x_A \in C$ and $x_A \in A$, and hence $x_A \in C \cap A$. Suppose that $y \in C \cap A$ so that $y \in C$ and hence there is a $B \in \mathcal{A}$ where $y = x_B$. We also have that $x_B = y \in A$. If $B \neq A$ then $x_B \in B$ and $x_B \in A$, which is not possible since B and A are disjoint as they are distinct elements of $\mathcal{A}$. So it must be that $B = A$ and hence $y = x_B = x_A$. Since y was arbitrary, this shows that $C \cap A$ has only a single element x_A. This suffices to show the choice axiom. $\square$

Exercise 19.10

Let A be a set; let $\{X_\alpha\}_{\alpha \in J}$ be an indexed family of spaces; and let $\{f_\alpha\}_{\alpha \in J}$ be an indexed family of functions $f_\alpha : A \to X_\alpha$.

(a) Show there is a unique coarsest topology $\mathcal{T}$ on A relative to which each of the functions f_α is continuous.

(b) Let

$$\mathcal{S}_\beta = \left\{ f_\beta^{-1}(U_\beta) \mid U_\beta \text{ is open in } X_\beta \right\},$$

and let $\mathcal{S} = \bigcup \mathcal{S}_\beta$. Show that $\mathcal{S}$ is a subbasis for $\mathcal{T}$.

(c) Show that a map $g : Y \to A$ is continuous relative to $\mathcal{T}$ if and only if each map $f_\alpha \circ g$ is continuous.

(d) Let $f : A \to \prod X_\alpha$ be defined by the equation

$$f(a) = (f_\alpha(a))_{\alpha \in J};$$

let Z denote the subspace $f(A)$ of the product space $\prod X_\alpha$. Show that the image under f of each element of $\mathcal{T}$ is an open set of Z.

Solution:

(a)

Proof. Let $\mathcal{C}$ be the collection of topologies on A relative to which each of the functions f_α is continuous. Clearly $\mathcal{C}$ is nonempty as the discrete topology is in $\mathcal{C}$ since every subset of A is open in it so that $f_\alpha(V_\alpha)$ is always open when V_α is open in X_α. Let $\mathcal{T} = \bigcap \mathcal{C}$, which is a topology on A by what was shown in Exercise 13.4a. We claim that this is the unique coarsest topology such that each f_α is continuous relative to it. To see this suppose that $\mathcal{T}'$ is any topology in such that each f_α is continuous relative to it, hence $\mathcal{T}' \in \mathcal{C}$. Then, for any open $U \in \mathcal{T} = \bigcap \mathcal{C}$ we of course have that $U \in \mathcal{T}'$ since $\mathcal{T}' \in \mathcal{C}$. Hence $\mathcal{T} \subset \mathcal{T}'$ since U was arbitrary so that $\mathcal{T}$ is courser than $\mathcal{T}'$, noting that it could of course be that $\mathcal{T} = \mathcal{T}'$ as well. Since $\mathcal{T}'$ was artbitary, this shows the desired result.

Of course it also must be that $\mathcal{T}$ is unique since, for any other $\mathcal{T}'$ that is a coarsest element of $\mathcal{C}$, we just showed above that $\mathcal{T} \subset \mathcal{T}'$ since $\mathcal{T}' \in \mathcal{C}$. But also $\mathcal{T} \supset \mathcal{T}'$ since $\mathcal{T}'$ must be coarser than $\mathcal{T}$ since $\mathcal{T} \in \mathcal{C}$. This shows that $\mathcal{T} = \mathcal{T}'$ so that $\mathcal{T}$ is unique since $\mathcal{T}'$ was arbitrary. This also follows from the more general fact that any smallest element in an order or partial order is always unique, and inclusion is always at least a partial order. $\square$

(b)

Proof. We show that $\mathcal{C}$ from part (a) is exactly the set of topologies on A that contain the subbasis $\mathcal{S}$. That is, we show that $\mathcal{T}' \in \mathcal{C}$ if and only if $\mathcal{S} \subset \mathcal{T}'$ when $\mathcal{T}'$ is a topology on A. Since the coarsest topology $\mathcal{T}$ from part (a) is defined as $\bigcap \mathcal{C}$, this shows that $\mathcal{T}$ is the topology generated from the subbasis $\mathcal{S}$ by Exercise 13.5.

($\Rightarrow$) Suppose that $\mathcal{T}' \in \mathcal{C}$ so that every f_α is continuous relative to $\mathcal{T}'$. Now consider any subbasis element $S \in \mathcal{S}$ so that $S = f_\beta^{-1}(U_\beta)$ for some $\beta \in J$ and some open set U_β in X_β. Then f_β is continuous relative to $\mathcal{T}'$ so that S is open with respect to $\mathcal{T}'$, and hence $S \in \mathcal{T}'$. This shows that $\mathcal{S} \subset \mathcal{T}'$ since S was arbitrary, hence $\mathcal{T}'$ contains $\mathcal{S}$.

($\Leftarrow$) Now suppose that $\mathcal{T}'$ is a topology on A that contains $\mathcal{S}$ so that $\mathcal{S} \subset \mathcal{T}'$. Consider any $\alpha \in J$ and any open set U_α of X_α. Then clearly $f_\alpha^{-1}(U_\alpha)$ is in $\mathcal{S}_\alpha$ so that it is also clearly in $\mathcal{S} = \bigcup \mathcal{S}_\beta$. Hence also $f_\alpha^{-1}(U_\alpha) \in \mathcal{T}'$ since $\mathcal{S} \subset \mathcal{T}'$. Therefore $f_\alpha^{-1}(U_\alpha)$ is open with respect to $\mathcal{T}'$, which shows that f_α is continuous relative to $\mathcal{T}'$ since U_α was an arbitrary open set of X_α. Since $\alpha \in J$ was also arbitrary, this shows that every f_α is continuous relative to $\mathcal{T}'$ so that $\mathcal{T}' \in \mathcal{C}$ by definition. $\square$

(c)

Proof. ($\Rightarrow$) Suppose that $g : Y \to A$ is continuous relative to $\mathcal{T}$. Consider any $\alpha \in J$ and any open set U_α of X_α. Then $f_\alpha^{-1}(U_\alpha)$ is open with respect to $\mathcal{T}$ since f_α is continuous relative to $\mathcal{T}$ since every f_α is. It then follows that $g^{-1}(f_\alpha^{-1}(U_\alpha))$ is open in Y since g is continuous relative to $\mathcal{T}$. From Exercise 2.4a we have that $g^{-1}(f_\alpha^{-1}(U_\alpha)) = (f_\alpha \circ g)^{-1}(U_\alpha)$, which shows that $f_\alpha \circ g$ is continuous since U_α was an arbitrary open set of X_α. Since $\alpha \in J$ was arbitrary, this shows the desired result.

($\Leftarrow$) Now suppose that every $f_\alpha \circ g$ is continuous and consider any open set U of A with respect to $\mathcal{T}$. Then by part (b) we have that U is an arbitrary union of finite intersections of subbasis elements $f_\alpha^{-1}(U_\alpha)$ for $\alpha \in J$ and open U_α in X_α. It then follows from Exercise 2.2 parts (b) and (c) that $g^{-1}(U)$ is an arbitrary union of finite intersections of sets $g^{-1}(f_\alpha^{-1}(U_\alpha))$. Again we have that each $g^{-1}(f_\alpha^{-1}(U_\alpha)) = (f_\alpha \circ g)^{-1}(U_\alpha)$ by Exercise 2.4a so that each of these sets is open in Y since every $f_\alpha \circ g$ is continuous. Hence $g^{-1}(U)$ is open as well since it is the arbitrary union of finite intersections of these open sets and Y is a topological space. Since U was an arbitrary open set of A with respect to $\mathcal{T}$, this shows that g is continuous relative to $\mathcal{T}$ as desired. $\square$

(d)

Proof. Suppose that U is any open set of A with respect to $\mathcal{T}$. Consider any $\mathbf{y} = (y_\alpha)_{\alpha \in J} \in f(U)$ so that there is an $a \in U$ where $f(a) = \mathbf{y}$. Since $a \in U$ and U is open in A, we have that there is a basis element B_A containing a where $B_A \subset U$. It then follows from part (b) that this basis element is a finite intersection of subbasis elements, hence $B_A = \bigcap_{\beta \in I} f_\beta^{-1}(U_\beta)$, where $I \subset J$ is finite and each U_β is open in X_β. Now define

$$V_\alpha = \begin{cases} U_\beta & \alpha \in I \\ X_\beta & \alpha \notin I \end{cases}$$

so that clearly the set $B_p = \prod V_\alpha$ is a basis element of $\prod X_\alpha$ in the product topology by Theorem 19.1 since I is finite. We then have that $B_Z = Z \cap B_p$ is a basis element of the subspace Z by Lemma 16.1.

Now, we have that $a \in U$ and $U \subset A$ so that $a \in A$ as well. It then follows that $\mathbf{y} = f(a) \in f(A) = Z$. For $\beta \in I$, we also have that $a \in f_\beta^{-1}(U_\beta)$ since the basis element $B_A = \bigcap_{\beta \in I} f_\beta^{-1}(U_\beta)$ contains a. Hence $f_\beta(a) \in U_\beta$. Since of course every other $f_\alpha(a) \in X_\alpha$ when $\alpha \notin I$, we have that $f_\alpha(a) \in V_\alpha$ for all $\alpha \in J$ and thus $\mathbf{y} = f(a) = (f_\alpha(a))_{\alpha \in J} \in \prod V_\alpha = B_p$. We therefore have that $\mathbf{y} \in Z \cap B_p = B_Z$ so that B_Z contains $\mathbf{y}$.

Lastly, consider any $\mathbf{z} = (z_\alpha)_{\alpha \in J} \in B_Z = Z \cap B_p$. Then $\mathbf{z} \in Z = f(A)$ so that there is an $x \in A$ where $f(x) = (f_\alpha(x))_{\alpha \in J} = \mathbf{z}$ and hence each $f_\alpha(x) = z_\alpha$. We also have that $\mathbf{z} \in B_p = \prod V_\alpha$ so that $z_\alpha \in V_\alpha$ for every $\alpha \in J$. In particular $f_\beta(x) = z_\beta \in V_\beta = U_\beta$ for all $\beta \in I$ so that $x \in f_\beta^{-1}(U_\beta)$. Therefore $x \in \bigcap_{\beta \in I} f_\beta^{-1}(U_\beta) = B_A$ so that also $x \in U$ since $B_A \subset U$. Then we have that $\mathbf{z} = f(x) \in f(U)$. Since $\mathbf{z}$ was arbitrary this shows that $B_Z \subset f(U)$.

We have thus shown that B_Z is a basis element of the subspace Z that contains $\mathbf{y}$ where $B_Z \subset f(U)$. Since $\mathbf{y}$ was an arbitrary element of $f(U)$, this suffices to show that $f(U)$ is open in the subspace Z as desired. $\square$