Analysis I (Tao) Notes

Steven Labalme

July 30, 2020

Contents

Chapter 1

Introduction

1.1 What Is Analysis?

6/14:

- **Analysis**: "The rigorous study of [mathematical] objects, with a focus on trying to pin down precisely and accurately the qualitative and quantitative behavior of those objects" (Tao, 2016, p. 1).
- **Real analysis**: "The analysis of the real numbers, sequences and series of real numbers, and real-valued functions" (Tao, 2016, p. 1).
- Real analysis is the theoretical foundation for **calculus**.
- **Calculus**: "The collection of computational algorithms which one uses to manipulate functions" (Tao, 2016, p. 1).
- Lists questions that can be answered with real analysis (motivation for studying it).

1.2 Why Do Analysis?

- Lists examples of contradictions in naïve calculus that must be resolved (and can be resolved with real analysis).

Chapter 2

Starting at the Beginning: The Natural Numbers

6/15:

- This text will begin by reviewing high school level material, but as rigorously as possible.
  - It will teach the skill of proving complicated properties from simpler ones, allowing you to understand why an "obvious" statement really is obvious.
  - One particularly important skill is the use of **mathematical induction**.
  - We will strive to eliminate **circularity**.

  Circularity: "Using an advanced fact to prove a more elementary fact, and then later using the elementary fact to prove the advanced fact" (Tao, 2016, p. 14).
- The number systems used in real analysis, listed in order of increasing sophistication, are the **naturals** $\mathbb{N}$[1], the **integers** $\mathbb{Z}$, the **rationals** $\mathbb{Q}$, and the **reals** $\mathbb{R}$.
  - **Complex numbers** $\mathbb{C}$ will only be used much later.
- This chapter will answer the question, "How does one actually *define* the natural numbers?"

2.1 The Peano Axioms

- **Peano Axioms**: First laid out by Guiseppe Peano, these are a standard way to define the natural numbers. They consist of Axioms 2.1-2.5, which follow.
  - From these five axioms and some from set theory, we can build all other number systems, create functions, and do algebra and calculus.
- How do you define operations on the naturals?
  - Complicated operations are defined in terms of simpler ones: Exponentiation is repeated multiplication, multiplication is repeated addition, and addition is repeated **incrementing**.
- **Incrementing**: The most fundamental operation — best thought of as counting forward by one number.
  - Incrementing is one of the fundamental concepts that allows us to define the natural numbers.
  - Let[2] $n{+}{+}$ denote the increment, or **successor**, of n.

[1] Note that in this text, the natural numbers will include 0. The natural numbers without 0 will be called the **positive integers** $\mathbb{Z}^+$.

[2] This notation is pulled from some computer languages such as *C*.

- For example, $3{+}{+} = 4$ and $(3{+}{+}){+}{+} = 5$.

- Let $x := y$ denote the statement, "x is defined to equal y."
- At this point, we can begin defining the natural numbers.

  Axiom 2.1. *0 is a natural number.*

  Axiom 2.2. *If n is a natural number, then $n{+}{+}$ is also a natural number.*

- To avoid having to use incrementation notation for every number, we adopt a convention.

  Definition 2.1. We define 1 to be the number $0{+}{+}$, 2 to be the number $(0{+}{+}){+}{+}$, 3 to be the number $((0{+}{+}){+}{+}){+}{+}$, etc.

- From these axioms, we can already prove things.

  Proposition 2.1. *3 is a natural number.*

  Proof. By Axiom 2.1, 0 is a natural number. By Axiom 2.2, $0{+}{+} = 1$ is a natural number. By Axiom 2.2 again, $1{+}{+} = 2$ is a natural number. By Axiom 2.2 again, $2{+}{+} = 3$ is a natural number. □

- It seems like Axioms 2.1 and 2.2 have us pretty well covered. However, what if the number system wraps around (e.g., if $3{+}{+} = 0$)? We can fix this with the following.

  Axiom 2.3. *0 is not the successor of any natural number; i.e., we have $n{+}{+} \neq 0$ for every natural number n.*

- We can now prove that $4 \neq 0$ (because $4 = 3{+}{+}$, $3 \in \mathbb{N}$, and $n{+}{+} \neq 0$).
- However, there are still issues — what if the number system hits a ceiling at 4, e.g., $4{+}{+} = 4$?
- A good way to prevent this kind of behavior is via the following.

  Axiom 2.4. *Different natural numbers must have different successors, i.e., if $n, m \in \mathbb{N}$ and $n \neq m$, then $n{+}{+} \neq m{+}{+}$. Equivalently*[3]*, if $n{+}{+} = m{+}{+}$, then $n = m$.*

- We can now prove propositions like the following, extending our anti-wrap around proving ability.

  Proposition 2.2. *6 is not equal to 2.*

  Proof. Suppose $6 = 2$. Then $5{+}{+} = 1{+}{+}$, so by Axiom 2.4, $5 = 1$. Then $4{+}{+} = 0{+}{+}$, so by Axiom 2.4, $4 = 0$, which contradicts our proof that $4 \neq 0$. □

6/16:

- Before going any further, we're going to need an **axiom schema**.
- **Axiom schema**: An axiom that functions as "a template for producing an (infinite) number of axioms, rather than a single axiom in its own right" (Tao, 2016, p. 20).

  Axiom 2.5 (Principle of mathematical induction)**.** *Let $P(n)$ be any property pertaining to a natural number n. Suppose that $P(0)$ is true, and suppose that whenever $P(n)$ is true, $P(n{+}{+})$ is also true. Then $P(n)$ is true for every natural number n.*

- Axiom 2.5 allows us to exclude numbers such as $0.5, 1.5, 2.5, \ldots$ from our number system because $P(n)$ is only true for $n \in 0, 1, 2, \ldots$

[3] This is an example of reformulating an implication using its **contrapositive** (see Section A.2). In the converse direction, it is the **axiom of substitution** (see Section A.7).

- Proposition 2.1.11 in Tao, 2016 is an excellent template for an induction proof.
- Note that there is only one natural number system — we could call $\{0, 1, 2, \dots\}$ and $\{O, I, II, III, \dots\}$ different number systems, but they are **isomorphic**, since a one-to-one correspondence exists between their elements and they obey the same rules.
- An interesting property of the naturals is that while every element is finite (0 is finite; if n is finite, then $n{+}{+}$ is finite), the set is infinite.
- In math, we define the natural numbers **axiomatically** as opposed to **constructively** — "we have not told you what the natural numbers are... we have only listed some things you can do with them... and some of the properties that they have" (Tao, 2016, p. 22).
  - This is the essence of treating objects **abstractly**, caring only about the properties of objects, not what they are or what they mean.
  - "The great discovery of the late nineteenth century was that numbers can be understood abstractly via axioms, without necessarily needing a concrete model; of course a mathematician can use any of these models [e.g., counting beads] when it is convenient, to aid his or her intuition and understanding, but they can also be just as easily discarded when they begin to get in the way [of understanding -3, $1/3$, $\sqrt{2}$, $3 + 4i$, ...]" (Tao, 2016, p. 23).
- With the axioms (and the concept of a function, which does not rely on said axioms), we can introduce recursive definitions, which will be useful in defining addition and multiplication.

  Proposition 2.3 (Recursive definitions)**.** *Suppose for each natural number n, we have some function $f_n : \mathbb{N} \to \mathbb{N}$ from the natural numbers to the natural numbers. Let c be a natural number. Then we can assign a unique natural number a_n to each natural number n, such that $a_0 = c$ and $a_{n++} = f_n(a_n)$ for each natural number n.*

  Proof. (Informal) We use induction. First, a single value c is given to a_0 (no other value $a_{n++} := f_n(a_n)$ will be assigned to 0 by Axiom 2.3). Given that a_n has a unique value, a_{n++} will have a unique value $f_n(a_n)$, distinct from any other a_{m++} by Axiom 2.4. □

2.2 Addition

- We can define addition recursively.

  Definition 2.2 (Addition of natural numbers)**.** Let m be a natural number. To add zero to m, we define $0 + m := m$. Now suppose inductively that we have defined how to add n to m. Then we can add $n{+}{+}$ to m by defining $(n{+}{+}) + m := (n + m){+}{+}$.
  - If we want to find $2 + 5$, we can find $0 + 5 = 5$, $1 + 5 = (0{+}{+}) + 5 = (0 + 5){+}{+} = 5{+}{+} = 6$, $2 + 5 = (1{+}{+}) + 5 = (1 + 5){+}{+} = 6{+}{+} = 7$.
- Let's now prove commutativity.

  Lemma 2.1. *For any natural number n, $n + 0 = n$.*

  Proof. Use induction. Since $0 + m = m$ for all $m \in \mathbb{N}$ and $0 \in \mathbb{N}$, $0 + 0 = 0$, proving the base case. If $n + 0 = n$, then $(n{+}{+}) + 0 = (n + 0){+}{+} = n{+}{+}$. This closes the induction. □

  Lemma 2.2. *For any $n, m \in \mathbb{N}$, $n + (m{+}{+}) = (n + m){+}{+}$.*

Proof. We keep m fixed and induct on n. Base case: if $n = 0$, then $0 + (m{+}{+}) = (m){+}{+} = (0 + m){+}{+}$. Induction step: if $n + (m{+}{+}) = (n + m){+}{+}$, then

$$\begin{aligned} (n{+}{+}) + (m{+}{+}) &= (n + (m{+}{+})){+}{+} && \text{Definition 2.2} \\ &= ((n + m){+}{+}){+}{+} && \text{Induction hypothesis} \\ &= ((n{+}{+}) + m){+}{+} && \text{Definition 2.2} \end{aligned}$$

This closes the induction. □

Proposition 2.4 (Addition is commutative)**.** *For any natural numbers n and m, $n + m = m + n$.*

Proof. For all $m \in \mathbb{N}$, Definition 2.2 gives us $0 + m = m$ and Lemma 2.1 gives us $m + 0 = m$. Since both of the previous statements equal m, $0 + m = m + 0$. Suppose inductively that $n \in \mathbb{N}$ and $n + m = m + n$. If this is true, then

$$\begin{aligned} (n{+}{+}) + m &= (n + m){+}{+} && \text{Definition 2.2} \\ &= (m + n){+}{+} && \text{Induction hypothesis} \\ &= m + (n{+}{+}) && \text{Lemma 2.2} \end{aligned}$$

This closes the induction. □

- And associativity (see Exercise 2.2.1).
- The next proposition deals with cancelling. Although we cannot use subtraction or negative numbers to prove it, it will be instrumental in allowing us to define subtraction and integers later.

  Proposition 2.5 (Cancellation law)**.** *Let a, b, c be natural numbers such that $a + b = a + c$. Then we have $b = c$.*

  Proof. We induct on a (keeping b, c fixed). Consider the base case $a = 0$. If $0 + b = 0 + c$ by assumption and $0 + b = b$ and $0 + c = c$ by Definition 2.2, then $b = c$. Suppose inductively that $a + b = a + c$ implies that $b = c$. We must prove that $(a{+}{+}) + b = (a{+}{+}) + c$ implies $b = c$. This may be done as follows.

$$\begin{aligned} (a{+}{+}) + b &= (a{+}{+}) + c && \text{Given} \\ (a + b){+}{+} &= (a + c){+}{+} && \text{Definition 2.2} \\ a + b &= a + c && \text{Axiom 2.4} \\ b &= c && \text{Induction hypothesis} \end{aligned}$$

  □

- **Positive natural numbers**: A natural number $n \neq 0$.

  Proposition 2.6. *If a is positive and b is a natural number, then $a + b$ is positive (and hence $b + a$ is also by Proposition 2.4).*

  Proof. We induct on b (keeping a fixed). In the base case, if $b = 0$, then $a + 0 = a$ (a positive number) by Lemma 2.1. Suppose inductively that $a + b$ is positive. Then $a + (b{+}{+}) = (a + b){+}{+}$ by Lemma 2.2, and $(a + b){+}{+}$ is positive by Axiom 2.3 — $a + (b{+}{+})$ is equal to the successor of a natural number, and the successor of a natural number is never 0, thus always positive. This closes the induction. □

  Corollary 2.1. *If $a, b \in \mathbb{N}$ and $a + b = 0$, then $a = 0$ and $b = 0$.*

  Proof. Suppose for the sake of contradiction that $a \neq 0$ or $b \neq 0$. If $a \neq 0$, then a is positive, and hence $a + b = 0$ is positive by Proposition 2.6, a contradiction. Similarly, if $b \neq 0$, then b is positive, and hence $a + b = 0$ is positive by Proposition 2.6, a contradiction. Thus, a and b must both be zero. □

- See Exercise 2.2.2 for another property of positive natural numbers.
- With addition, we can begin to order the natural numbers.

  Definition 2.3 (Ordering of the natural numbers)**.** Let $n, m \in \mathbb{N}$. We say that n is **greater than or equal to** m and write $n \geq m$ or $m \leq n$ iff we have $n = m + a$ for some $a \in \mathbb{N}$. We say that n is **strictly greater than** m and write $n > m$ or $m < n$ iff $n \geq m$ and $n \neq m$.

- Note that $n{+}{+} > n$ for any n.
- See Exercise 2.2.3 for more on ordering.
- We can now prove the trichotomy.

  Proposition 2.7 (Trichotomy of order for natural numbers)**.** *Let a and b be natural numbers. Then exactly one of the following statements is true: $a < b$, $a = b$, or $a > b$.*

  Proof. See Exercise 2.2.4 to fill in the gaps.

  First, show that no two (or three) of the statements can hold simultaneously. If $a < b$ or $a > b$, then $a \neq b$ by definition. Also, if $a > b$ and $a < b$, then $a = b$, a contradiction.

  Second, show that at least one of the statements is always true. We induct on a (keeping b fixed). When $a = 0$, we have $0 \leq b$ for all b (see Exercise 2.2.4a), so we either have $0 = b$ or $0 < b$, which proves the base case. Now suppose inductively that we have proven the proposition for a. From the trichotomy of a, there are three cases: $a < b$, $a = b$, and $a > b$. If $a > b$, then $a{+}{+} > b$ (see Exercise 2.2.4b). If $a = b$, then $a{+}{+} > b$ (see Exercise 2.2.4c). If $a < b$, then $a{+}{+} \leq b$ by Proposition 2.9. Thus, either $a{+}{+} = b$ or $a{+}{+} < b$. This closes the induction. □

6/17:
- With order, we can obtain a stronger version of induction (see Exercise 2.2.5).
  - Strong induction is usually used with $m_0 = 0$ or $m_0 = 1$.

Exercises

6/16:
1. Prove the following proposition. Hint: fix two of the variables and induct on the third.

   Proposition 2.8 (Addition is associative)**.** *For any natural numbers a, b, c, we have $(a + b) + c = a + (b + c)$.*

   Proof. We first need a lemma.

   Lemma 2.3. *The sum of two natural numbers $n + m$ is a natural number.*

   Proof. We induct on n (keeping m fixed). By Axiom 2.1, $0 \in \mathbb{N}$. Since $m \in \mathbb{N}$, by Definition 2.2, $0 + m$ (the sum of two natural numbers) equals m (a natural number). Thus, the base case holds. Suppose inductively that $n + m$ is a natural number. Then $(n{+}{+}) + m = (n + m){+}{+}$ by Definition 2.2, $n + m$ is a natural number by the induction hypothesis, and $(n + m){+}{+}$ is a natural number by Axiom 2.2. This closes the induction. □

   Now we induct on a (keeping b, c fixed). By the lemma, $b + c$ is a natural number and can be treated as such. Consider the base case $a = 0$. In this case, $0 + (b + c) = b + c$ and $0 + b = b$ by Definition 2.2, so $0 + (b + c) = b + c = (0 + b) + c$. Now suppose inductively that $a + (b + c) = (a + b) + c$. Then

   $$\begin{aligned} (a{+}{+}) + (b + c) &= (a + (b + c)){+}{+} && \text{Definition 2.2} \\ &= ((a + b) + c){+}{+} && \text{Induction hypothesis} \\ &= ((a + b){+}{+}) + c && \text{Definition 2.2} \\ &= ((a{+}{+}) + b) + c && \text{Definition 2.2} \end{aligned}$$

   This closes the induction. □

2. Prove the following lemma. Hint: use induction.

   Lemma 2.4. *Let a be a positive number. Then there exists exactly one natural number b such that $b{+}{+} = a$.*

   Proof. We induct on a. Consider the base case $a = 1$. $1 = 0{+}{+}$ by definition, and by Axiom 2.4, 0 is the only b satisfying $1 = b{+}{+}$. Now suppose inductively that a has only one b satisfying $b{+}{+} = a$. Then $a{+}{+}$ has only one natural number (namely a) satisfying $a{+}{+} = a{+}{+}$. This closes the induction. □

3. Prove the following proposition. Hint: you will need many of the preceding propositions, corollaries, and lemmas.

   Proposition 2.9 (Basic properties of order for natural numbers)**.** *Let a, b, c be natural numbers. Then*

   (*a*) (*Order is reflexive*) $a \geq a$.

   Proof. By Lemma 2.1, $a = a + 0$. The previous expression is in the form $n = m + a$; thus, by Definition 2.3, $a \geq a$. □

   (*b*) (*Order is transitive*) *If $a \geq b$ and $b \geq c$, then $a \geq c$.*

   Proof. If $a \geq b$ and $b \geq c$, then $a = b + n$ and $b = c + m$, respectively, for some $n, m \in \mathbb{N}$. Substituting, $a = (c + m) + n$. By Proposition 2.8, $a = c + (m + n)$. By Lemma 2.3, $m + n$ is a natural number. The previous expression is in the form $n = m + a$; thus, by Definition 2.3, $a \geq c$. □

   (*c*) (*Order is anti-symmetric*) *If $a \geq b$ and $b \geq a$, then $a = b$.*

   Proof. If $a \geq b$ and $b \geq a$, then $a = b + n$ and $b = a + m$, respectively, for some $n, m \in \mathbb{N}$. Substituting, $a = (a + m) + n$. By Proposition 2.8, $a = a + (m + n)$. By Lemma 2.1, $a + 0 = a + (m + n)$. By Proposition 2.5, $0 = m + n$. By Corollary 2.1, m and n both equal 0. Thus, $a = b + 0 = b$ (or $b = a + 0 = a$) by Lemma 2.1. □

   (*d*) (*Addition preserves order*) *$a \geq b$ iff $a + c \geq b + c$.*

   Proof. If $a + c \geq b + c$, then $a + c = (b + c) + n$ for some $n \in \mathbb{N}$. Then

   $$\begin{aligned} c + a &= n + (b + c) && \text{Proposition 2.4} \\ c + a &= (n + b) + c && \text{Proposition 2.8} \\ c + a &= c + (n + b) && \text{Proposition 2.4} \\ a &= n + b && \text{Proposition 2.5} \\ a &= b + n && \text{Proposition 2.4} \end{aligned}$$

   Thus, $a \geq b$. □

   (*e*) *$a < b$ iff $a{+}{+} \leq b$.*

   Proof. If $a{+}{+} \leq b$, then $b = (a{+}{+}) + n$ for some $n \in \mathbb{N}$. Then

   $$\begin{aligned} b &= (a + n){+}{+} && \text{Definition 2.2} \\ &= a + (n{+}{+}) && \text{Lemma 2.2} \end{aligned}$$

   Since $n{+}{+}$ is a natural number (Axiom 2.2), the above proves that $a \leq b$. By Axiom 2.3, $n{+}{+} \neq 0$. Thus, $b \neq a$ (suppose for the sake of contradiction that $b = a$. Then $b = b + 0 = a + (n{+}{+})$ implies by Proposition 2.5 that $0 = n{+}{+}$, a contradiction). By definition, since $a \leq b$ and $b \neq a$, $a < b$. □

   (*f*) *$a < b$ iff $b = a + d$ for some positive number d.*

Proof. As a positive number, d is a natural number by definition. Thus, $b = a + d$ implies $a \leq b$. Since d is a positive number, $d \neq 0$. For the reasons outlined in the previous proof, this implies that $b \neq a$. Thus, $a < b$. □

4. Justify the three statements marked (why?) in the proof of Proposition 2.7.

   (a) *If n is a natural number, then $0 \leq n$.*

   Proof. We induct on n. By Proposition 2.9, $0 \geq 0$, proving the base case. Suppose inductively that $n \geq 0$. We know that $n{+}{+} \geq n$ (since $n{+}{+} = (n + 0){+}{+} = n + 0{+}{+}$), so by Proposition 2.9, $n{+}{+} \geq n$ and $n \geq 0$ transitively imply $n{+}{+} \geq 0$. □

   (b) *Let a, b be natural numbers. Then if $a > b$, $a{+}{+} > b$.*

   Proof. We first need a lemma.

   Lemma 2.5. *If $a > b$ and $b > c$, then $a > c$.*

   Proof. If $a > b$ and $b > c$, then $a = b + n$ and $b = c + m$, respectively, for some positive numbers n, m. Substituting, $a = (c + m) + n$. By Proposition 2.8, $a = c + (m + n)$. By Proposition 2.6, $m + n$ is a positive number. Thus, by Proposition 2.9, $a > c$. □

   $a{+}{+} > a$. By the lemma, $a{+}{+} > a$ and $a > b$ imply that $a{+}{+} > b$. □

   (c) *Let a, b be natural numbers. Then if $a = b$, $a{+}{+} > b$.*

   Proof. $a{+}{+} > a$. Since $a = b$, substituting gives $a{+}{+} > b$. □

6/17: 5. Prove the following proposition. Hint: define $Q(n)$ to be the property that $P(m)$ is true for all $m_0 \leq m < n$; note that $Q(n)$ is vacuously true when $n < m_0$.

Proposition 2.10 (Strong principle of induction)**.** *Let m_0 be a natural number, and let $P(m)$ be a property pertaining to an arbitrary natural number m. Suppose that for each $m \geq m_0$, we have the following implication: if $P(m')$ is true for all natural numbers $m_0 \leq m' < m$, then $P(m)$ is also true. (In particular, this means that $P(m_0)$ is true, since in this case the hypothesis is vacuous.) Then we can conclude that $P(m)$ is true for all natural numbers $m \geq m_0$.*

Proof. Let n be a natural number satisfying $n \geq m_0$ and let $Q(n)$ be the property that $P(m)$ is true for all $m_0 \leq m < n$. We induct on n.

For the base case $n = 0$, we want to show that $Q(0)$ is true. Since $0 \leq m_0$ (Exercise 2.2.4a), either $0 = m_0$ or $0 < m_0$ (Proposition 2.7). We treat these cases separately. If $0 = m_0$, then $Q(0)$ is vacuously true (since there is no $m \in \mathbb{N}$ such that $0 \leq m < 0$, of course $P(m)$ is true for all $0 \leq m < 0$ [because there are no cases, $P(m)$ is true in all cases; it's also false, but that's besides the point.]) If $0 < m_0$, then $Q(0)$ is also vacuously true (since there is no $m \in \mathbb{N}$ such that $0 < m_0 \leq m < 0$).

Suppose inductively that for some $n \geq m_0$, $Q(n)$ is true, i.e., "$P(m)$ is true for all $m_0 \leq m < n$" is true. We want to show that $Q(n{+}{+})$ is true. By the definition of P in the hypothesis, $Q(n)$ is true implies $P(n)$ is true. Thus, $P(m)$ is true for all $m_0 \leq m < n{+}{+}$, so $Q(n{+}{+})$ is true. This closes the induction.

Since $Q(n)$ is true for all $n \geq m_0$, $P(n)$ is true for all $n \geq m_0$ (proof modified from Ojo, 2019). □

6. Let n be a natural number, and let $P(m)$ be a property pertaining to the natural numbers such that whenever $P(m{+}{+})$ is true, then $P(m)$ is true. Suppose that $P(n)$ is also true. Prove that $P(m)$ is true for all natural numbers $m \leq n$; this is known as the **principle of backwards induction**. Hint: apply induction to the variable n.

Proof. We induct on n. For the base case $n = 0$, we want to show that $P(m)$ is true for all natural numbers $m \leq n$. Since we are supposing that $P(n) = P(0)$ is true and 0 is the only number $m \in \mathbb{N}$ satisfying $m \leq 0$, the base case is true.

Suppose inductively that we have proved for $P(n)$ true that $P(m)$ is true for all natural numbers $m \leq n$. We want to show that for $P(n{+}{+})$ true, $P(m)$ is true for all natural numbers $m \leq n{+}{+}$. If $P(n{+}{+})$ is true, then clearly $P(m)$ is true for all natural numbers $m = n$. By definition, $P(n{+}{+})$ true implies $P(n)$ true, and $P(n)$ true implies that $P(m)$ is true for all natural numbers $m \leq n$. Thus, $P(m)$ is true for all natural numbers $m \leq n$ or $m = n{+}{+}$; these two statements can be combined into $m \leq n{+}{+}$. This closes the induction. □

2.3 Multiplication

6/18:

- At this point, all properties of addition can be used without supplying a justification.
- The definition of multiplication is very similar to the definition of addition.

  Definition 2.4 (Multiplication of natural numbers)**.** Let m be a natural number. To multiply 0 to m, we define $0 \times m := 0$. Now suppose inductively that we have defined how to multiply n to m. Then we can multiply $n{+}{+}$ to m by defining $(n{+}{+}) \times m := (n \times m) + m$.

- See Exercise 2.3.1 for a proof of the commutativity of multiplication.
- At this point, we abbreviate $n \times m$ as nm and use the convention that multiplication takes precedence over addition (i.e., $ab + c = (a \times b) + c$).
- See Exercise 2.3.2 for a proof of the zero product property.
- We can now prove the distributive law.

  Proposition 2.11 (Distributive law)**.** *For any natural numbers a, b, c, we have $a(b+c) = ab+ac$ and $(b+c)a = ba+ca$.*

  Proof[4]*.* Since multiplication is commutative, we need only show the first identity $a(b+c) = ab+ac$. We keep a and b fixed, and use induction on c. Let's prove the base case $c = 0$, i.e., $a(b+0) = ab+a0$. The left-hand side is ab, while the right-hand side is $ab + 0 = ab$, so we are done with the base case. Now let us suppose inductively that $a(b+c) = ab+ac$, and let us prove that $a(b+(c{+}{+})) = ab+a(c{+}{+})$. The left-hand side is $a((b+c){+}{+}) = a(b+c)+a$, while the right-hand side is $ab+ac+a = a(b+c)+a$ by the induction hypothesis, and so we can close the induction. □

- See Exercise 2.3.3 for a proof of the associativity of multiplication.
- Bringing back order, we can prove that multiplication preserves order.

  Proposition 2.12 (Multiplication preserves order)**.** *If a, b are natural numbers such that $a < b$, and c is positive, then $ac < bc$.*

  Proof. Since $a < b$, we have $b = a + d$ for some positive d. Multiplying by c and using the distributive law, we obtain $bc = ac + dc$. Since c, d are positive, dc is positive, and hence $ac < bc$ as desired. □

- Although we still haven't introduced division, we can introduce a multiplicative analogy of the cancellation law.

  Corollary 2.2 (Cancellation law)**.** *Let a, b, c be natural numbers such that $ac = bc$ and c is non-zero. Then $a = b$.*

[4]Note that this could be accomplished by induction on a. However, for the sake of introducing new concepts, I will transcribe the alternative method from Tao, 2016.

Proof[5]. By the Proposition 2.7, we have three cases: $a < b$, $a = b$, or $a > b$. Suppose first that $a < b$, then by Proposition 2.12, we have $ac < bc$, which contradicts our hypothesis that $ac = bc$. We can obtain a similar contradiction when $a > b$. Thus, the only possibility is that $a = b$, as desired. □

- At this point, we can deduce all of the rules of algebra involving addition and multiplication (see Exercise 2.3.4 for such an example).
- With addition and multiplication now defined, we will rarely see incrementation moving forward. Regardless, we can always use $n{+}{+} = n + 1$ to describe it (this is not circular since $n + 1$ was originally defined from $n{+}{+}$).
- See Exercise 2.3.5 for a proof of the Euclidean algorithm.
  - This algorithm marks the beginning of **number theory**, which is important but will not be covered any further in this text.
- The definition of exponentiation is very similar to the definitions of addition and multiplication.

  Definition 2.5 (Exponentiation of natural numbers)**.** Let m be a natural number. To raise m to the power 0, we define $m^0 := 1$; in particular, we define $0^0 := 1$. Now suppose recursively that m^n has been defined for some natural number n. Then we define $m^{n++} := m^n \times m$.

- Exponentiation will not be explored too deeply here — wait until after we define the integers and rationals.

Exercises

1. Prove the following lemma. Hint: modify the proofs of Lemmas 2.1 and 2.2 and Proposition 2.4.

   Lemma 2.6 (Multiplication is commutative)**.** *Let n, m be natural numbers. Then $n \times m = m \times n$.*

   Proof. We first need two lemmas.

   Lemma 2.7. *For any natural number n, $n \times 0 = 0$.*

   Proof. We induct on n. For the base case $n = 0$, Definition 2.4 gives us $0 \times 0 = 0$. Suppose inductively that $n \times 0 = 0$. Then $(n{+}{+}) \times 0 = (n \times 0) + 0$ by Definition 2.4, which equals $0 + 0$ by the induction hypothesis, which equals 0. This closes the induction. □

   Lemma 2.8. *Let n, m be natural numbers. Then $n \times (m{+}{+}) = (n \times m) + n$.*

   Proof. We induct on n (keeping m fixed). For the base case $n = 0$, Definition 2.4 gives us $0 \times (m{+}{+}) = 0$ and $(0 \times m) + 0 = 0$. Thus, $0 \times (m{+}{+}) = (0 \times m) + 0$, proving the base case. Suppose inductively that $n \times (m{+}{+}) = (n \times m) + n$. Then

   $$\begin{aligned} (n{+}{+}) \times (m{+}{+}) &= (n \times (m{+}{+})) + m && \text{Definition 2.4} \\ &= ((n \times m) + n) + m && \text{Induction hypothesis} \\ &= ((n \times m) + m) + n \\ &= ((n{+}{+}) \times m) + n && \text{Definition 2.4} \end{aligned}$$

   This closes the induction. □

[5] This could be accomplished by induction on c.

Now for the primary proof, we induct on n (keeping m fixed). For the base case $n = 0$, Definition 2.4 gives us $0 \times m = 0$ while Lemma 2.7 gives us $m \times 0 = 0$. Thus, $0 \times m = m \times 0$, proving the base case. Now suppose inductively that $n \times m = m \times n$. Then

$$\begin{aligned} (n{+}{+}) \times m &= (n \times m) + m && \text{Definition 2.4} \\ &= (m \times n) + m && \text{Induction hypothesis} \\ &= m \times (n{+}{+}) && \text{Lemma 2.8} \end{aligned}$$

This closes the induction. □

2. Prove the following lemma. Hint: prove the second statement first.

Lemma 2.9 (Positive natural numbers have no zero divisors)**.** *Let n, m be natural numbers. Then $n \times m = 0$ if and only if at least one of n, m is equal to zero. In particular, if n and m are both positive, then nm is also positive.*

Proof. We begin by proving the second statement. We induct on n (keeping m fixed). For the base case, n is actually equal to 1 (since 1 is the smallest positive number). We want to show that $1m$ is positive. By Definition 2.4, $1m = 0m + m = 0 + m = m$, which is positive by assumption. Suppose inductively that nm is positive. Then $(n{+}{+})m = nm + m$. Since $nm + m$ is the sum of two positive numbers, it is positive (Proposition 2.6). This closes the induction.

Suppose for the sake of contradiction that n, m are both positive natural numbers and $nm = 0$. Since $nm = 0$, it is not positive by definition. But this contradicts the previous assertion that nm must be positive if n, m are both positive. Thus, at least one of n, m is not positive, implying by definition that at least one of n, m is equal to 0. □

3. Prove the following proposition. Hint: modify the proof of Proposition 2.8 and use the distributive law.

Proposition 2.13 (Multiplication is associative)**.** *For any natural numbers a, b, c, we have $(a \times b) \times c = a \times (b \times c)$.*

Proof. We first need a lemma.

Lemma 2.10. *For any natural numbers n, m, the product nm is a natural number.*

Proof. We induct on n (keeping m fixed). For the base case $n = 0$, $0m = 0$ by Definition 2.4, which is a natural number. Now suppose inductively that nm is a natural number. Then $(n{+}{+})m = nm + m$, which is the sum of two natural numbers. By Lemma 2.3, $nm + m$ is a natural number. This closes the induction. □

Now for the primary proof, we induct on a (keeping b, c fixed). For the base case $a = 0$, $(0 \times b) \times c = 0 \times c = 0$ by two applications of Definition 2.4. Also, $0 \times (b \times c) = 0$ by Definition 2.4 and the lemma (which asserts that bc can be treated as a natural number). Thus, $(0 \times b) \times c = 0 \times (b \times c)$ transitively, proving the base case. Now suppose inductively that $(a \times b) \times c = a \times (b \times c)$. Then

$$\begin{aligned} ((a{+}{+}) \times b) \times c &= ((a \times b) + b) \times c && \text{Definition 2.4} \\ &= (a \times b) \times c + b \times c && \text{Proposition 2.11} \\ &= a \times (b \times c) + b \times c && \text{Induction hypothesis} \\ &= (a{+}{+}) \times (b \times c) && \text{Definition 2.4} \end{aligned}$$

This closes the induction. □

4. Prove the identity $(a + b)^2 = a^2 + 2ab + b^2$ for all natural numbers a, b.

Proof. We first need two lemmas.

Lemma 2.11. *For any natural numbers n, m, $nm = \underbrace{m + \cdots + m}_{n \text{ times}}$.*

Proof. We induct on n (keeping m fixed). For the base case $n = 0$, $0m = 0$ by Definition 2.4. 0 is equal to the sum of zero m's, proving the base case. Suppose inductively that

$$nm = \underbrace{m + \cdots + m}_{n \text{ times}}$$

Then

$$(n{+}{+})m = nm + m = \underbrace{m + \cdots + m}_{n \text{ times}} + m = \underbrace{m + \cdots + m}_{n{+}{+} \text{ times}}$$

This closes the induction. □

Lemma 2.12. *For any natural numbers n, m, $m^n = \underbrace{m \times \cdots \times m}_{n \text{ times}}$.*

Proof. We induct on n (keeping m fixed). For the base case $n = 0$, $m^0 = 1$ by Definition 2.5. 1 is equal to the product of zero m's, proving the base case. Suppose inductively that

$$m^n = \underbrace{m \times \cdots \times m}_{n \text{ times}}$$

Then

$$m^{n{+}{+}} = m^n \times m = \underbrace{m \times \cdots \times m}_{n \text{ times}} \times m = \underbrace{m \times \cdots \times m}_{n{+}{+} \text{ times}}$$

This closes the induction. □

Now for the primary proof:

$$\begin{aligned}
(a+b)^2 &= (a+b)(a+b) && \text{Lemma 2.12} \\
&= (a+b)a + (a+b)b && \text{Proposition 2.11} \\
&= aa + ba + ab + bb && \text{Proposition 2.11} \\
&= a^2 + ba + ab + b^2 && \text{Lemma 2.12} \\
&= a^2 + (ab + ab) + b^2 && \text{Lemma 2.6} \\
&= a^2 + 2ab + b^2 && \text{Lemma 2.11}
\end{aligned}$$

□

5. Prove the following proposition. Hint: fix q and induct on n.

Proposition 2.14 (Euclidean algorithm)**.** *Let n be a natural number, and let q be a positive number. Then there exist natural numbers m, r such that $0 \leq r < q$ and $n = mq + r$.*

Proof. We induct on n (keeping q fixed). For the base case $n = 0$, choose $m = 0$ and $r = 0$. In both cases, 0 is a natural number, and in the latter case, 0 satisfies $0 \leq 0 < q$ since $0 \leq 0$ by Proposition 2.9 and $0 < q$ for all positive q. Since m, r meet all necessary conditions, $0q + 0 = 0$ by Definition 2.4 and 2.2, and $n = 0$ by hypothesis, the base case holds.

Suppose inductively that $n = mq + r$. We want to show that $n{+}{+} = m'q + r'$ for some m', r' satisfying the conditions that m and r must satisfy, respectively. To begin, we can show an equality between m, r and m', r' as follows.

$$\begin{aligned} m'q + r' &= n{+}{+} \\ &= n + 1 \\ &= mq + r + 1 \end{aligned}$$

From this, it would seem logical to choose $m' := m$ and $r' := r + 1$. However, since $r < q$, $r + 1 \leq q$, so by Proposition 2.7 we have two cases to consider. If $r + 1 < q$, it will indeed suffice to choose $m' := m$ and $r' := r + 1$. Both of these choices satisfy all of the necessary requirements. If $r + 1 = q$, we must choose differently. Choose $m' := m + 1$ and $r' := 0$. It may be more difficult to understand these choices, but they do work: $m + 1$ and 0 are both natural numbers, 0 satisfies $0 \leq 0 < q$ as discussed earlier, and the following shows that the choices maintain the equality with $n{+}{+}$.

$$\begin{aligned} (m + 1)q + 0 &= mq + q \\ &= mq + r + 1 \\ &= n + 1 \\ &= n{+}{+} \end{aligned}$$

With both cases treated, the induction may be closed. □

Chapter 3

Set Theory

7/4:

- Set theory will be frequently used in the subsequent chapters; it is part of the foundation of almost every other branch of mathematics.
  - Note that Euclidean geometry will not be defined — we will use the Cartesian coordinate system's parallel with the real numbers instead.
- This chapter covers the elementary aspects, Chapter **??** covers more advanced topics, and the finer subtleties are well beyond the scope of this text.

3.1 Fundamentals

- We define sets axiomatically, as we did with the natural numbers[1].

  Axiom 3.1 (Sets are objects)**.** *If A is a set, then A is also an object. In particular, given two sets A and B, it is meaningful to ask whether A is also an element of B.*

- Note that while all sets are objects, not all objects are sets.
  - For example, 1 is not a set while $\{1\}$ is.
  - Note, though, that **pure set theory** considers all objects to be sets. However, impure set theory (where some objects are not sets) is conceptually easier to deal with.
    - Since both types are equal for the purposes of mathematics, we will take a middle-ground approach.
- If x, y are objects and A a set, then the statement $x \in A$ is either true or false. Note that $x \in y$ is neither true nor false, simply meaningless.
- We now define equality for sets.

  Definition 3.1 (Equality of sets)**.** Two sets A and B are equal, $A = B$, iff every element of A is an element of B and vice versa. To put it another way, $A = B$ if and only if every element x of A belongs also to B, and every element y of B belongs also to A.

- Note that this implies that repetition of elements does not effect equality ($\{3, 3\} = \{3\}$, for example).
- It can be proven that this notion of equality is reflexive, symmetric, and transitive (see Exercise 3.1.1).

7/14:

- Since $x \in A$ and $A = B$ implies $x \in B$, the $\in$ relation obeys the axiom of substitution as well.

[1] Note that the following list of axioms will be somewhat overcomplete, as some axioms may be derived from others. However, this is helpful for pedagogical reasons, and there is no real harm being done.

- Thus, any operation defined in terms of the $\in$ relation obeys the axiom of substitution.

- We define sets in an analogous way to how we defined natural numbers from 0, onward.

  Axiom 3.2 (Empty set)**.** *There exists a set $\emptyset$ (also denoted $\{\}$), known as the empty set, which contains no elements, i.e., for every object x, we have $x \notin \emptyset$.*

  - Note that there is only one empty set — if $\emptyset$ and $\emptyset'$ were supposedly distinct empty sets, then Definition 3.1 would prove that $\emptyset = \emptyset'$.

- **Non-empty set**: A set that "is not equal to the empty set" (Tao, 2016, p. 36).

- Non-empty sets must contain at least one object.

  Lemma 3.1 (Single choice)**.** *Let A be a non-empty set. Then there exists an object x such that $x \in A$.*

  Proof. Suppose for the sake of contradiction that no object x exists such that $x \in A$, i.e., for all objects x, we have $x \notin A$. By Axiom 3.2, we have $x \notin \emptyset$ either. Thus, $x \in A \iff x \in \emptyset$ (denoting logical equivalence; both statements are equally false), so, by Definition 3.1, $A = \emptyset$, a contradiction. □

- There exist more sets than just the empty set.

  Axiom 3.3 (Singleton sets and pair sets)**.** *If a is an object, then there exists a set $\{a\}$ whose only element is a, i.e., for every object y, we have $y \in \{a\}$ if and only if $y = a$; we refer to $\{a\}$ as the **singleton set** whose element is a. Furthermore, if a and b are objects, then there exists a set $\{a, b\}$ whose elements are a and b; i.e., for every object y, we have $y \in \{a, b\}$ if and only if $y = a$ or $y = b$; we refer to this set as the **pair set** formed by a and b.*

- By Definition 3.1, there exists only one singleton set for each object a and only one pair set for any two objects a, b.

- Note that the singleton set axiom follows from the pair set axiom, and the pair set axiom follows from the singleton set axiom and the pairwise union axiom, below.

- As alluded to, the pairwise union axiom allows us to build sets with more than two elements.

  Axiom 3.4 (Pairwise union)**.** *Given any two sets A, B, there exists a set $A \cup B$ called the **union** $A \cup B$ of A and B, whose elements consist of all the elements which belong to A or B or both. In other words, for any object x,*

  $$x \in A \cup B \iff (x \in A \text{ or } x \in B)$$

- The $\cup$ operation obeys the axiom of substitution (if $A = A'$, then $A \cup B = A' \cup B$).

- We now prove some basic properties of unions (one below and three in Exercise 3.1.3).

  Lemma 3.2. *If A, B, C are sets, then the union operation is associative, i.e., $(A \cup B) \cup C = A \cup (B \cup C)$.*

  Proof. By Definition 3.1, showing that every element x of $(A \cup B) \cup C$ is an element of $A \cup (B \cup C)$ and vice versa will suffice to prove this lemma. Suppose first that $x \in (A \cup B) \cup C$. By Axiom 3.4, this means that at least one of $x \in A \cup B$ or $x \in C$ is true. We now divide into two cases. If $x \in C$, then by Axiom 3.4, $x \in B \cup C$, and, so, by Axiom 3.4 again, we have $x \in A \cup (B \cup C)$. Now suppose instead that $x \in A \cup B$. Then by Axiom 3.4, $x \in A$ or $x \in B$. If $x \in A$, then $x \in A \cup (B \cup C)$ by Axiom 3.4, while if $x \in B$, then by consecutive applications of Axiom 3.4, we have $x \in B \cup C$ and, hence, $x \in A \cup (B \cup C)$. A similar argument shows that every element of $A \cup (B \cup C)$ lies in $(A \cup B) \cup C$, and so $(A \cup B) \cup C = A \cup (B \cup C)$, as desired. □

- As a consequence of the above, we are free to write $A \cup B \cup C \cup \cdots$ to denote repeated unions without having to use parentheses.

- We can also now define triplet sets ($\{a,b,c\} := \{a\} \cup \{b\} \cup \{c\}$), quadruplet sets, and so forth.
  - However, we cannot yet define a set of n objects or an infinite set.
- Note that addition and union are analogous, but importantly *not* identical.
- Some sets are "larger" than others; hence, subsets.

  Definition 3.2 (Subsets)**.** Let A, B be sets. We say that A is a **subset** of B, denoted $A \subseteq B$, iff every element of A is also an element of B, i.e., for any object x, $x \in A \implies x \in B$. We say that A is a **proper subset** of B, denoted $A \subsetneq B$ if $A \subseteq B$ and $A \neq B$.

- The $\subseteq$ and $\subsetneq$ operations obey the axiom of substitution (since both $=$ and $\in$, the two component operations of $\subseteq$ and $\subsetneq$, obey it).
- Note that $A \subseteq A$ and $\emptyset \subseteq A$ for any set A.
- Note that less than or equal to and subset are analogous, but not identical, either (see below for one related property and Exercise 3.1.4 for two more).

  Proposition 3.1 (Sets are partially ordered by set inclusion 1)**.** *Let A, B, C be sets. If $A \subseteq B$ and $B \subseteq C$, then $A \subseteq C$.*

  Proof. Suppose that $A \subseteq B$ and $B \subseteq C$. To prove that $A \subseteq C$, we have to prove that every element of A is an element of C. Let $x \in A$. Then $x \in B$ (since $A \subseteq B$), implying that $x \in C$ (since $B \subseteq C$). $\square$

- There exist relations between subsets and unions (see Exercise 3.1.7).
- Note this difference between $\subsetneq$ and $<$: Given any two distinct natural numbers n, m, one is smaller than the other (Proposition 2.7). However, given any two distinct sets, it is not in general true that one is a subset of the other. This is why we say that sets are **partially ordered** while the natural numbers (for example) are **totally ordered** (see Definitions **??** and **??**, respectively).
- Note that $\in$ and $\subseteq$ are distinct ($2 \in \{1,2,3\}$, but $2 \not\subseteq \{1,2,3\}$; similarly, $\{2\} \subseteq \{1,2,3\}$, but $\{2\} \notin \{1,2,3\}$).
- It is important to distinguish sets from their elements, for they can have different properties ($\mathbb{N}$ is an infinite set of finite elements, and $\{\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}\}$ is a finite set of infinite objects).
- We now formally state that it is acceptable to create subsets.

  Axiom 3.5 (Axiom of specification)**.** *Let A be a set, and for each $x \in A$, let $P(x)$ be a property pertaining to x (i.e., $P(x)$ is either a true statement or a false statement). Then there exists a set, called $\{x \in A : P(x)$ is true$\}$ (or simply $\{x \in A : P(x)\}$ for short), whose elements are precisely the elements x in A for which $P(x)$ is true. In other words, for any object y,*

  $$y \in \{x \in A : P(x) \text{ is true}\} \iff (y \in A \text{ and } P(y) \text{ is true})$$

- *Also known as* **axiom of separation**.
- Specification obeys the axiom of substitution (since $\in$ obeys it).
- Sometimes $\{x \in A : P(x)\}$ is denoted by $\{x \in A \,|\, P(x)\}$ (useful when we need the colon for something else, e.g., $f : X \to Y$).
- We use Axiom 3.5 to define intersections.

Definition 3.3 (Intersections)**.** The **intersection** $S_1 \cap S_2$ of two sets is defined to be the set

$$S_1 \cap S_2 := \{x \in S_1 : x \in S_2\}$$

In other words, $S_1 \cap S_2$ consists of all the elements which belong to both S_1 and S_2. Thus, for all objects x,

$$x \in S_1 \cap S_2 \iff x \in S_1 \text{ and } x \in S_2$$

- The $\cap$ operation obeys the axiom of substitution (since $\in$ obeys it).
  - Note that since $\cap$ is defined in terms of more primitive operations, it is well-defined.
- Problems with the English word, "and."
  - It can mean union or intersection depending on the context.
    - If X, Y are sets, "the set of elements of X and elements of Y" refers to $X \cup Y$, e.g., "the set of singles and males."
    - If X, Y are sets, "the set of objects that are elements of X and elements of Y" refers to $X \cap Y$, e.g., "the set of people who are single and male."
  - It can also denote addition.
    - "2 and 3 is 5" means $2 + 3 = 5$.
  - "One reason we resort to mathematical symbols instead of English words such as 'and' is that mathematical symbols always have a precise and unambiguous meaning, whereas one must often look very carefully at the context in order to work out what an English word means" (Tao, 2016, p. 42).
- **Disjoint** (sets): Two sets A, B such that $A \cap B = \emptyset$.
- **Distinct** (sets): Two sets A, B such that $A \neq B$.
  - Note that $\emptyset$ and $\emptyset$ are disjoint but not distinct.
- We also use Axiom 3.5 to define difference sets.

  Definition 3.4 (Difference sets)**.** Given two sets A and B, we define the set $A - B$ or $A \setminus B$ to be the set A with any elements of B removed, i.e.,

$$A \setminus B := \{x \in A : x \notin B\}$$

- For example, $\{1, 2, 3, 4\} \setminus \{2, 4, 6\} = \{1, 3\}$ — in many cases, $B \subseteq A$, but not necessarily.

7/15:

- See Exercise 3.1.6 for some basic properties of unions, intersections, and difference sets.
  - The de Morgan laws are named after the logician Augustus De Morgan (1806-1871).
  - The **laws of Boolean algebra** (the contents of Proposition 3.4) are named after the mathematician George Boole (1815-1864).
    - They are applicable to a number of objects other than sets (e.g., laws of propositional logic).
  - Note the **duality** in Proposition 3.4, manifesting itself in the certain symmetry between $\cup$ and $\cap$, and X and $\emptyset$.
- **Duality**: "Two distinct properties or objects being dual to each other" (Tao, 2016, p. 43).
- We can do a lot, but we still wish to do more. For example, we'd like to transform sets (say, take $\{3, 5, 9\}$ and increment each object to yield $\{4, 6, 10\}$).

Axiom 3.6 (Replacement)**.** *Let A be a set. For any object $x \in A$, and any object y, suppose we have a statement $P(x, y)$ pertaining to x and y such that for each $x \in A$, there is at most one y for which $P(x, y)$ is true. Then there exists a set $\{y : P(x, y)$ is true for some $x \in A\}$, such that for any object z,*

$$z \in \{y : P(x, y) \text{ is true for some } x \in A\} \iff P(x, z) \text{ is true for some } x \in A$$

- For example, let $A := \{3, 5, 9\}$ and let $P(x, y)$ be the statement $y = x{+}{+}$. By Axiom 2.4, for every $x \in A$, there is exactly one y for which $P(x, y)$ is true (namely, the successor of x). Thus, by Axiom 3.6, the set $\{y : y = x{+}{+}$ for some $x \in \{3, 5, 9\}\}$ exists. It is clearly the same set as $\{4, 6, 10\}$.
- The set obtained can be smaller than the original set, e.g., $\{y : y = 1$ for some $x \in \{3, 5, 9\}\} = \{1\}$.
- We often abbreviate the set specified in Axiom 3.6 to one of the following.

$$\{y : y = f(x) \text{ for some } x \in A\}$$
$$\{f(x) : x \in A\}$$
$$\{f(x) \,\big|\, x \in A\}$$

- We can combine Axioms 3.6 and 3.5, i.e., $\{f(x) : x \in A; P(x)$ is true$\}$, e.g., $\{n{+}{+} : n \in \{3, 5, 9\}; n < 6\} = \{4, 6\}$.
- Although we have assumed that natural numbers are objects in several examples up to this point, we must formalize this notion.

  Axiom 3.7 (Infinity)**.** *There exists a set $\mathbb{N}$, whose elements are called the natural numbers, as well as an object 0 in $\mathbb{N}$, and an object $n{+}{+}$ assigned to every natural number $n \in N$, such that the Peano axioms hold.*

- Axiom 3.7 is called the **axiom of infinity** because "it introduces the most basic example of an infinite set, namely the set of natural numbers $\mathbb{N}$" (Tao, 2016, p. 44).

Exercises

7/4: 1. Show that the definition of equality in Definition 3.1 is reflexive, symmetric, and transitive[2].

Proof. Given a set A, suppose $A \neq A$. Then, by Definition 3.1, every element of A is not an element of A, a contradiction. Thus, $A = A$.

Let sets $A = B$. Then, by Definition 3.1, every element x of A belongs also to B, and every element y of B belongs also to A. Identically, every element y of B belongs also to A, and every element X of A belongs also to B. Thus, $B = A$.

Let sets $A = B$ and $B = C$. Then, by Definition 3.1, every element x of A belongs also to B, and every element y of B belongs also to A. Similarly, every element y of B belongs also to C, and every element z of C belongs also to B. Since $x \in A \Rightarrow x \in B \Rightarrow x \in C$, and $y \in C \Rightarrow y \in B \Rightarrow y \in A$, $A = C$. □

7/14: 2. Using only Definition 3.1, Axiom 3.1, Axiom 3.2, and Axiom 3.3, prove that the sets $\emptyset$, $\{\emptyset\}$, $\{\{\emptyset\}\}$, and $\{\emptyset, \{\emptyset\}\}$ are all distinct (i.e., no two of them are equal to each other).

Proof. First, we show that all sets are distinct from the empty set. Axiom 3.1 asserts that $\emptyset$ and $\{\emptyset\}$ are objects. By Axiom 3.3, we have $\emptyset \in \{\emptyset\}$, $\emptyset \in \{\emptyset, \{\emptyset\}\}$, and $\{\emptyset\} \in \{\{\emptyset\}\}$. Since $x \notin \emptyset$ for all objects x (Axiom 3.2), $\{\emptyset\}$, $\{\{\emptyset\}\}$, and $\{\emptyset, \{\emptyset\}\}$ all contain objects that $\emptyset$ does not (namely, $\emptyset$, $\{\emptyset\}$, and $\emptyset$, respectively). Thus, by Definition 3.1, $\emptyset \neq \{\emptyset\}$, $\emptyset \neq \{\{\emptyset\}\}$, and $\emptyset \neq \{\emptyset, \{\emptyset\}\}$.

[2]Note that since Definition 3.1 should be an axiom (should axiomatize equality and all that that entails for sets), this exercise is silly (see Tao, 2020).

Next, we show that $\{\emptyset\} \neq \{\emptyset, \{\emptyset\}\}$. By Axiom 3.3, $\{\emptyset\} \in \{\emptyset, \{\emptyset\}\}$ and $y \in \{\emptyset\}$ iff $y = \emptyset$. Since $\{\emptyset\} \neq \emptyset$ (see above), $\{\emptyset\} \notin \{\emptyset\}$. Thus, $\{\emptyset, \{\emptyset\}\}$ contains an object that $\{\emptyset\}$ does not, implying by Definition 3.1 that $\{\emptyset\} \neq \{\emptyset, \{\emptyset\}\}$.

Lastly, we show that $\{\emptyset\} \neq \{\{\emptyset\}\}$ and that $\{\emptyset, \{\emptyset\}\} \neq \{\{\emptyset\}\}$. We proceed in a similar manner to the above. By Axiom 3.3, $\emptyset \in \{\emptyset\}$, $\emptyset \in \{\emptyset, \{\emptyset\}\}$, and $y \in \{\{\emptyset\}\}$ iff $y = \{\emptyset\}$. Since $\emptyset \neq \{\emptyset\}$ (see above), $\emptyset \notin \{\{\emptyset\}\}$. Thus, $\{\emptyset\}$ and $\{\emptyset, \{\emptyset\}\}$ both contain an object that $\{\{\emptyset\}\}$ does not, implying by Definition 3.1 that $\{\emptyset\} \neq \{\{\emptyset\}\}$ and that $\{\emptyset, \{\emptyset\}\} \neq \{\{\emptyset\}\}$.

□

3. Prove the following lemmas.

Lemma 3.3. *If a and b are objects, then $\{a, b\} = \{a\} \cup \{b\}$.*

Proof. By Definition 3.1, it will suffice to show that every element x of $\{a, b\}$ is an element of $\{a\} \cup \{b\}$ and vice versa. By Axiom 3.3, if $x \in \{a, b\}$, then $x = a$ or $x = b$. By Axiom 3.4, if $x \in \{a\} \cup \{b\}$, then $x \in \{a\}$ or $x \in \{b\}$, implying by Axiom 3.3 that $x = a$ or $x = b$. Thus, the elements of $\{a, b\}$ and of $\{a\} \cup \{b\}$ are both a, b, so by Definition 3.1, the sets are equal. □

Lemma 3.4. *If A, B, C are sets, then the union operation is commutative, i.e., $A \cup B = B \cup A$.*

Proof. By Definition 3.1, it will suffice to show that every element x of $A \cup B$ is an element of $B \cup A$ and vice versa. Let $x \in A \cup B$. By Axiom 3.4, $x \in A$ or $x \in B$. If, on the one hand, $x \in A$, then $x \in B \cup A$ (by Axiom 3.4). If, on the other hand, $x \in B$, then $x \in B \cup A$ (by Axiom 3.4). A similar argument holds if we choose an element $y \in B \cup A$ first. □

Lemma 3.5. *If A is a set, then $A = A \cup \emptyset = \emptyset \cup A = A \cup A$.*

Proof. First, we show that $A \cup \emptyset = \emptyset \cup A$. This is a direct consequence of Lemma 3.3.

Next, we show that $A = A \cup \emptyset$. By Definition 3.1, it will suffice to show that every element x of A is an element of $A \cup \emptyset$ and vice versa. By Axiom 3.4, every element x of A is an element of $A \cup \emptyset$. Now let $x \in A \cup \emptyset$. Then by Axiom 3.4, $x \in A$ or $x \in \emptyset$. By Axiom 3.2, $x \notin \emptyset$, so $x \in A$. Thus, every element of $A \cup \emptyset$ is an element of A. We now have by the transitive property that $A = A \cup \emptyset = \emptyset \cup A$.

Lastly, we show that $A = A \cup A$. By Definition 3.1, it will suffice to show that every element x of A is an element of $A \cup A$ and vice versa. By Axiom 3.4, every element x of A is an element of $A \cup A$. Now let $x \in A \cup A$. Then by Axiom 3.4, $x \in A$ or $x \in A$, implying $x \in A$. We have, at last, by the transitive property that $A = A \cup \emptyset = \emptyset \cup A = A \cup A$. □

4. Prove the following propositions.

Proposition 3.2 (Sets are partially ordered by set inclusion 2)**.** *Let A, B be sets. If $A \subseteq B$ and $B \subseteq A$, then $A = B$.*

Proof. Suppose that $A \subseteq B$ and $B \subseteq A$. By Definition 3.2, $A \subseteq B$ implies that every element of A is also an element of B and $B \subseteq A$ implies that every element of B is also an element of A. Thus, by Definition 3.1, $A = B$. □

Proposition 3.3 (Sets are partially ordered by set inclusion 3)**.** *Let A, B, C be sets. If $A \subsetneq B$ and $B \subsetneq C$, then $A \subsetneq C$.*

Proof. Suppose that $A \subsetneq B$ and $B \subsetneq C$. Then, by Definition 3.2, $A \subseteq B$ and $B \subseteq C$, implying $A \subseteq C$ by the first claim proved. Since $A \subsetneq B$, $A \neq B$ (implying, by Definition 3.1, that every element of A is not an element of B or every element of B is not an element of A) and every element of A is an element of B; hence, every element of B is not an element of A. Therefore, B *must* contain some element that A does not. Similarly, $B \subsetneq C$ implies that C must contain some element that B does not. Hence, C contains at least two elements that A does not, proving that $A \neq C$, too. □

5. Let A, B be sets. Show that the three statements $A \subseteq B$, $A \cup B = B$, and $A \cap B = A$ are logically equivalent (any one of them implies the other two).

Proof. First, we show that $A \subseteq B \implies (A \cup B = B$ and $A \cap B = A)$. Next, we show that $A \cup B = B \implies A \subseteq B$ (which, in turn, implies $A \cap B = A$). Lastly, we show that $A \cap B = A \implies A \subseteq B$ (which, in turn, implies $A \cup B = B$). Let's begin.

Suppose that $A \subseteq B$. To prove $A \cup B = B$, Definition 3.1 tells us that it will suffice to show that every element x of $A \cup B$ is an element of B and vice versa. By Axiom 3.4, $x \in B \implies x \in A \cup B$. By Axiom 3.4, $x \in A \cup B \implies (x \in A$ or $x \in B)$. By Definition 3.2, $A \subseteq B$ means that $x \in A \implies x \in B$. Thus, $x \in A \cup B \implies (x \in A$ or $x \in B) \implies x \in B$. Therefore, if $A \subseteq B$, then $A \cup B = B$. To prove that $A \cap B = A$, Definition 3.1 tells us that it will suffice to show that every element x of $A \cap B$ is an element of A and vice versa. By Definition 3.3, $x \in A \cap B \implies x \in A$ (and $x \in B$). Since $x \in A \implies x \in B$ (see above), $x \in A \implies (x \in A$ and $x \in B) \implies x \in A \cap B$ (Definition 3.3). Therefore, if $A \subseteq B$, then $A \cap B = A$.

Suppose that $A \cup B = B$. To prove $A \subseteq B$, Definition 3.2 tells us that it will suffice to show that every element of A is also an element of B. By Axiom 3.4, $x \in A \implies x \in A \cup B$. By Definition 3.1, $y \in A \cup B \implies y \in B$. Thus, $x \in A \implies x \in A \cup B \implies x \in B$. Therefore, if $A \cup B = B$, $A \subseteq B$ (and $A \cap B = A$).

Suppose that $A \cap B = A$. To prove that $A \subseteq B$, Definition 3.2 tells us that it will suffice to show that every element of A is also an element of B. By Definition 3.1, $x \in A \implies x \in A \cap B$. By Definition 3.3, $y \in A \cap B \implies y \in B$. Thus, $x \in A \implies x \in A \cap B \implies x \in B$. Therefore, if $A \cap B = A$, $A \subseteq B$ (and $A \cup B = B$). □

7/15: 6. Prove the following proposition. (Hint: one can use some of these claims to prove others. Some of the claims have also appeared previously in Lemma 3.2 and Exercise 3.1.3.)

Proposition 3.4 (Sets form a boolean algebra)**.** *Let A, B, C be sets, and let X be a set containing A, B, C as subsets.*

(a) *(Minimal element) We have $A \cup \emptyset = A$ and $A \cap \emptyset = \emptyset$.*

Proof. See Exercise 3.1.3 for the first claim.

To prove $A \cap \emptyset = \emptyset$, Definition 3.1 tells us that it will suffice to show that every element x of $A \cap \emptyset$ is an element of $\emptyset$ and vice versa. First off, "every element $x \in \emptyset$ is an element of $A \cap \emptyset$" is vacuously true (by Axiom 3.2, there exists no $x \in \emptyset$). In the other direction, suppose for the sake of contradiction that $x \in A \cap \emptyset$ for some object x. By Definition 3.3, $x \in A$ and $x \in \emptyset$. But $x \in \emptyset$ contradicts Axiom 3.2. Therefore, $x \notin A \cap \emptyset$ for all objects x. Thus, the statement "every element $x \in A \cap \emptyset$ is an element of $\emptyset$" is vacuously true (there exists no $x \in A \cap \emptyset$). □

(b) *(Maximal element) We have $A \cup X = X$ and $A \cap X = A$.*

Proof. See Exercise 3.1.5. □

(c) *(Identity) We have $A \cap A = A$ and $A \cup A = A$.*

Proof. To prove that $A \cap A = A$, Definition 3.1 tells us that it will suffice to show that every element x of $A \cap A$ is an element of A and vice versa. By Definition 3.3, $x \in A \cap A \implies (x \in A$ and $x \in A) \implies x \in A$. On the other hand, $x \in A \implies (x \in A$ and $x \in A)$ (idempotent law for conjunction), $(x \in A$ and $x \in A) \implies x \in \{x \in A : x \in A\}$ (Axiom 3.5), and $x \in \{x \in A : x \in A\} \implies x \in A \cap A$ (Definition 3.3).

See Exercise 3.1.3 for the second claim. □

(d) *(Commutativity) We have $A \cup B = B \cup A$ and $A \cap B = B \cap A$.*

Proof. See Exercise 3.1.3 for the first claim.

To prove $A \cap B = B \cap A$, Definition 3.1 tells us that it will suffice to show that every element x of $A \cap B$ is an element of $B \cap A$ and vice versa. By two applications of Definition 3.3 with the commutative law for conjunction in between, $x \in A \cap B \implies (x \in A \text{ and } x \in B) \implies (x \in B \text{ and } x \in A) \implies x \in B \cap A$. A similar argument works in the opposite direction. □

(e) *(Associativity) We have* $(A \cup B) \cup C = A \cup (B \cup C)$ *and* $(A \cap B) \cap C = A \cap (B \cap C)$.

Proof. See Lemma 3.2 for the first claim.

By Definition 3.1, showing that every element x of $(A \cap B) \cap C$ is an element of $A \cap (B \cap C)$ and vice versa will suffice to prove this lemma. Suppose first that $x \in (A \cap B) \cap C$. By Definition 3.3, $x \in A \cap B$ and $x \in C$, which implies by a second application of Definition 3.3 that $x \in A$ and $x \in B$ and $x \in C$. It follows by consecutive applications of Definition 3.3 that $x \in A$ and $x \in B \cap C$, and that $x \in A \cap (B \cap C)$. A similar argument shows that every element of $A \cap (B \cap C)$ lies in $(A \cap B) \cap C$. □

(f) *(Distributivity) We have* $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ *and* $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.

Proof. To prove $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$, Definition 3.1 tells us that it will suffice to show that every element x of $A \cap (B \cup C)$ is an element of $(A \cap B) \cup (A \cap C)$ and vice versa. By Definition 3.3, $x \in A \cap (B \cup C) \implies (x \in A \text{ and } x \in B \cup C)$. By Axiom 3.4, $x \in B \cup C \implies x \in B$ or $x \in C$. Thus, we know that $x \in A$, and we know that $x \in B$ or $x \in C$ (or both). We divide into two cases. Suppose first that $x \in B$. Since $x \in A$ as well, $x \in A \cap B$ (Definition 3.3). This implies by Axiom 3.4 that $x \in (A \cap B) \cup (A \cap C)$. Now suppose that $x \in C$. Since $x \in A$ as well, $x \in A \cap C$ (Definition 3.3). This implies by Axiom 3.4 that $x \in (A \cap B) \cup (A \cap C)$. A similar argument shows that every element of $(A \cap B) \cup (A \cap C)$ lies in $A \cap (B \cup C)$.

The second proof is similar to the first (and similar to all the rest of the proofs written for this exercise thus far). Thus, for the sake of variety, we will do this one entirely symbolically, using the laws of propositional logic from Section A.4.

$$
\begin{aligned}
x \in A \cup (B \cap C) &\implies (x \in A) \vee (x \in B \cap C) && \text{Axiom 3.4} \\
&\implies (x \in A) \vee ((x \in B) \wedge (x \in C)) && \text{Definition 3.3} \\
&\implies ((x \in A) \wedge (x \in B)) \vee ((x \in A) \wedge (x \in C)) && \text{Distributive law for disjunction} \\
&\implies (x \in A \cap B) \vee (x \in A \cap C) && \text{Definition 3.3} \\
&\implies x \in (A \cap B) \cup (A \cap C) && \text{Axiom 3.4}
\end{aligned}
$$

A similar argument works in reverse. □

(g) *(Partition) We have* $A \cup (X \setminus A) = X$ *and* $A \cap (X \setminus A) = \emptyset$.

Proof. To prove $A \cup (X \setminus A) = X$, Definition 3.1 tells us that it will suffice to show that every element x of $A \cup (X \setminus A)$ is an element of X and vice versa. Suppose $x \in A \cup (X \setminus A)$. Then by Axiom 3.4, $x \in A$ or $x \in X \setminus A$. We divide into two cases. Suppose first that $x \in A$. Since $A \subseteq X$, Definition 3.2 asserts that $x \in X$. Now suppose that $x \in X \setminus A$. This implies by Definition 3.4 that $x \in \{y \in X : y \notin A\}$, which means by Axiom 3.5 that $x \in X$ (and $x \notin A$, but that's not relevant). On the other hand, suppose $x \in X$. Naturally, either $x \in A$ or $x \notin A$ ($x \in A$ is false). If $x \in A$, then $x \in A \cup (X \setminus A)$ (Axiom 3.4). If $x \notin A$, since $x \in X$ as well, Axiom 3.5 asserts that $x \in \{x \in X : x \notin A\}$, which, by Definition 3.4, implies that $x \in X \setminus A$. This, in turn, implies that $x \in A \cup (X \setminus A)$ (Axiom 3.4).

To prove $A \cap (X \setminus A) = \emptyset$, it suffices to prove that for every object x, we have $x \notin A \cap (X \setminus A)$ (because of the uniqueness of the empty set). Suppose for the sake of contradiction that $x \in A \cap (X \setminus A)$ for some object x. By Definition 3.3, $x \in A$ and $x \in X \setminus A$. By Definition 3.4, $x \in \{y \in X : y \notin A\}$. By Axiom 3.5, $x \in X$ and $x \notin A$. But this contradicts the previously derived fact that $x \in A$. Therefore, $x \notin A \cap (X \setminus A)$ for all objects x. □

(h) *(De Morgan laws) We have* $X \setminus (A \cup B) = (X \setminus A) \cap (X \setminus B)$ *and* $X \setminus (A \cap B) = (X \setminus A) \cup (X \setminus B)$.

Proof. To prove $X \setminus (A \cup B) = (X \setminus A) \cap (X \setminus B)$, Definition 3.1 tells us that it will suffice to show that every element x of $X \setminus (A \cup B)$ is an element of $(X \setminus A) \cap (X \setminus B)$ and vice versa. Suppose that $x \in X \setminus (A \cup B)$. Then by Definition 3.4, $x \in \{y \in X : y \notin A \cup B\}$. By Axiom 3.5, $x \in X$ and $x \notin A \cup B$. By the inverse of Axiom 3.4 (which is a valid assertion since Axiom 3.4 asserts the logical equivalence of "$x \in A \cup B$" and "$x \in A$ and $x \in B$"), $x \notin A \cup B \implies (x \notin A$ and $x \notin B)$. By Axiom 3.5 and Definition 3.4, $(x \in X$ and $x \notin A) \implies x \in \{y \in X : y \notin A\} \implies x \in X \setminus A$. Similarly, $(x \in X$ and $x \notin B) \implies x \in \{y \in X : y \notin B\} \implies x \in X \setminus B$. Thus, by Definition 3.3, $x \in (X \setminus A) \cap (X \setminus B)$. A similar, reversed argument will work in the other direction.

To prove $X \setminus (A \cap B) = (X \setminus A) \cup (X \setminus B)$, Definition 3.1 tells us that it will suffice to show that every element x of $X \setminus (A \cap B)$ is an element of $(X \setminus A) \cup (X \setminus B)$ and vice versa. By Definition 3.4 and Axiom 3.5, $x \in X \setminus (A \cap B) \implies x \in \{y \in X : y \notin A \cap B\} \implies (x \in X$ and $x \notin A \cap B)$. By the inverse of Definition 3.3 (which, again, is a valid assertion since $x \in A \cap B \leftrightarrow ((x \in A) \wedge (x \in B))$), $x \notin A \cap B \implies (x \notin A$ or $x \notin B)$. We divide into two cases. Suppose $x \notin A$. Then by Axiom 3.5 and Definition 3.4, $(x \in X$ and $x \notin A) \implies x \in \{y \in X : y \notin A\} \implies x \in X \setminus A$. Thus, by Axiom 3.4, $x \in (X \setminus A) \cup (X \setminus B)$. Now suppose $x \notin B$. Then by Axiom 3.5 and Definition 3.4, $(x \in X$ and $x \notin B) \implies x \in \{y \in X : y \notin B\} \implies x \in X \setminus B$. Thus, by Axiom 3.4, $x \in (X \setminus A) \cup (X \setminus B)$. □

7/17: 7. Let A, B, C be sets. Show that $A \cap B \subseteq A$ and $A \cap B \subseteq B$. Furthermore, show that $C \subseteq A$ and $C \subseteq B$ iff $C \subseteq A \cap B$. In a similar spirit, show that $A \subseteq A \cup B$ and $B \subseteq A \cup B$, and furthermore that $A \subseteq C$ and $B \subseteq C$ iff $A \cup B \subseteq C$.

Proof. To prove that $A \cap B \subseteq A$ and $A \cap B \subseteq B$, Definition 3.2 tells us that it will suffice to show that every element of $A \cap B$ is an element of A and B. By Definition 3.3, $x \in A \cap B \implies (x \in A$ and $x \in B)$.

7/19: To prove that $C \subseteq A$ and $C \subseteq B$ iff $C \subseteq A \cap B$, it will suffice to show that $C \subseteq A$ and $C \subseteq B$ imply $C \subseteq A \cap B$ and vice versa. Suppose first that $C \subseteq A$ and $C \subseteq B$. Then by two applications of Definition 3.2, $x \in C \implies x \in A$ and $x \in C \implies x \in B$. By Definition 3.3, $(x \in A$ and $x \in B) \implies x \in A \cap B$. Thus, every element x of C is an element of $A \cap B$, so by Definition 3.2, $C \subseteq A \cap B$. Now suppose that $C \subseteq A \cap B$. Then by Definition 3.2, $x \in C \implies x \in A \cap B$. By Definition 3.3, $x \in A \cap B \implies (x \in A$ and $x \in B)$. Thus, every element x of C is an element of A and B, so by two applications of Definition 3.2, $C \subseteq A$ and $C \subseteq B$.

To prove that $A \subseteq A \cup B$ and that $B \subseteq A \cup B$, Definition 3.2 tells us that it will suffice to show that every element x of A is an element of $A \cup B$ and that every element y of B is an element of $A \cup B$, respectively. By two applications of Axiom 3.4, $x \in A \implies x \in A \cup B$, and $y \in B \implies y \in A \cup B$.

To prove that $A \subseteq C$ and $B \subseteq C$ iff $A \cup B \subseteq C$, it will suffice to show that $A \subseteq C$ and $B \subseteq C$ imply $A \cup B \subseteq C$ and vice versa. Suppose first that $A \subseteq C$ and $B \subseteq C$. By Axiom 3.4, $x \in A \cup B \implies (x \in A$ or $x \in B)$. We divide into two cases. If $x \in A$, then Definition 3.2 ensures that $x \in C$. If $x \in B$, then Definition 3.2 similarly ensures that $x \in C$. Thus, either way, $x \in A \cup B \implies x \in C$, so by Definition 3.2, $A \cup B \subseteq C$. Now suppose that $A \cup B \subseteq C$. By Axiom 3.4 followed by Definition 3.2, $x \in A \implies x \in A \cup B \implies x \in C$. A similar argument shows that $x \in B \implies x \in C$. Thus, $A \subseteq C$ and $B \subseteq C$. □

8. Let A, B be sets. Prove the **absorption laws** $A \cap (A \cup B) = A$ and $A \cup (A \cap B) = A$.

Proof. By Exercise 3.1.7, $A \subseteq A \cup B$. By Exercise 3.1.6b (with $A = A$ and $X = A \cup B$), $A \cap (A \cup B) = A$. By Exercise 3.1.7, $A \cap B \subseteq A$. By Exercise 3.1.6b (with $A = A \cap B$ and $X = A$), $A \cup (A \cap B) = A$. □

9. Let A, B, X be sets such that $A \cup B = X$ and $A \cap B = \emptyset$. Show that $A = X \setminus B$ and $B = X \setminus A$.

Proof. To prove $A = X \setminus B$, Definition 3.1 tells us that it will suffice to show that every element x of A is an element of $X \setminus B$ and vice versa. Suppose first that $x \in A$. Then by Axiom 3.4, $x \in A \cup B$. Since $A \cup B = X$, we have by Definition 3.1 that $x \in X$. Now suppose for the sake of contradiction

that $x \in B$. Since $x \in A$ as well, by Definition 3.3, $x \in A \cap B$. But $A \cap B = \emptyset$, so by Definition 3.1, $x \in \emptyset$, which contradicts Axiom 3.2. Therefore, $x \notin B$. Since $x \in X$ and $x \notin B$, by Definition 3.4, $x \in X \setminus B$. Now suppose that $x \in X \setminus B$. Then by Definition 3.4, $x \in X$ and $x \notin B$. By Definition 3.1, $x \in A \cup B$. Thus, by Axiom 3.4, $x \in A$ or $x \in B$. Since $x \notin B$, x must be an element of A. A similar argument shows that $B = X \setminus A$. □

7/22: 10. Let A and B be sets. Show that the three sets $A \setminus B$, $A \cap B$, and $B \setminus A$ are disjoint, and that their union is $A \cup B$.

Proof. To show that $A \setminus B$, $A \cap B$, and $B \setminus A$ are disjoint, we must show that

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

We may do this by showing that no object x is an element of any of the left three sets above (because of the uniqueness of the empty set). We do this via three contradiction proofs, as follows.

Suppose for the sake of contradiction that $x \in (A \setminus B) \cap (A \cap B)$. Then by Definition 3.3, $x \in A \setminus B$ and $x \in A \cap B$. Since $x \in A \setminus B$, Definition 3.4 tells us that $x \notin B$ (and $x \in A$). But since $x \in A \cap B$, Definition 3.3 tells us that $x \in B$ (and $x \in A$), a contradiction.

A similar argument to the above can handle $(A \cap B) \cap (B \setminus A)$.

Suppose for the sake of contradiction that $x \in (A \setminus B) \cap (B \setminus A)$. Then by Definition 3.3, $x \in A \setminus B$ and $x \in B \setminus A$. By the first statement, $x \in A$ and $x \notin B$, while by the second statement, $x \in B$ and $x \notin A$, two contradictions.

We now turn our attention to proving the following.

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

We can actually prove this solely on the basis of prior results (and one additional lemma).

Lemma 3.6. *Let A and B be sets. Show that $(A \cup B) \setminus A = B \setminus A$.*

Proof. To prove $(A \cup B) \setminus A = B \setminus A$, Definition 3.1 tells us that it will suffice to show that every element x of $(A \cup B) \setminus A$ is an element of $B \setminus A$ and vice versa. Suppose first that $x \in (A \cup B) \setminus A$. Then by Definition 3.4, $x \in A \cup B$ and $x \notin A$. Since $x \in A \cup B$, by Axiom 3.4, $x \in A$ or $x \in B$. But $x \notin A$, so x must be an element of B. Having established that $x \in B$ and $x \notin A$, Definition 3.4 tells us that $x \in B \setminus A$. Now suppose that $x \in B \setminus A$. Then by Definition 3.4, $x \in B$ and $x \notin A$. Since $x \in B$, by Axiom 3.4, $x \in A \cup B$. Consequently, by Definition 3.4, $x \in (A \cup B) \setminus A$. □

Now we can begin. By Exercise 3.1.7, $A \cap B \subseteq A$ and $A \subseteq A \cup B$. This implies by Proposition 3.1 that $A \cap B \subseteq A \cup B$. Thus,

$$\begin{aligned} A \cup B &= (A \cap B) \cup ((A \cup B) \setminus (A \cap B)) && \text{Exercise 3.1.6g} \\ &= (A \cap B) \cup ((A \cup B) \setminus A) \cup ((A \cup B) \setminus B) && \text{Exercise 3.1.6h} \\ &= (A \cap B) \cup (B \setminus A) \cup (A \setminus B) && \text{Lemma 3.6} \end{aligned}$$

□

11. Show that the axiom of replacement implies the axiom of specification.

Proof. Let $P(x, y)$ be the statement "$y = x$ and $P(y)$ is true." Then by Axiom 3.6, there exists a set

$$\begin{aligned} \{y : P(x, y) \text{ is true for some } x \in A\} &= \{y : y = x \text{ and } P(y) \text{ is true for some } x \in A\} \\ &= \{y : y \in A \text{ and } P(y) \text{ is true}\} \\ &= \{y \in A : P(y) \text{ is true}\} \end{aligned}$$

Moreover,

$$\begin{aligned} z \in \{y \in A : P(y) \text{ is true}\} &\implies z \in \{y : P(x, y) \text{ is true for some } x \in A\} \\ &\implies P(x, z) \text{ is true for some } x \in A \\ &\implies z = x \text{ and } P(z) \text{ is true for some } x \in A \\ &\implies z \in A \text{ and } P(z) \text{ is true} \end{aligned}$$

The above logic also works in reverse. Thus, all the tenets of Axiom 3.5 have been shown to follow from Axiom 3.6 (proof modified from Jiang, 2020). □

3.2 Russell's Paradox

- Suppose that we could unify the multitude of axioms in Section 3.1 into a single axiom. The following would be a good candidate (in fact, it implies the majority of the Section 3.1 axioms — see Exercise 3.2.1).

  Axiom 3.8 (Universal specification)**.** *(Dangerous!) Suppose for every object x we have a property $P(x)$ pertaining to x (so that for every x, $P(x)$ is either a true statement or a false statement). Then there exists a set $\{x : P(x)$ is true$\}$ such that for every object y,*

  $$y \in \{x : P(x) \text{ is true}\} \iff P(y) \text{ is true}$$

- *Also known as* **axiom of comprehension**.
- Basically, Axiom 3.8 asserts that "every property corresponds to a set" (Tao, 2016, p. 46).
- Unfortunately, Axiom 3.8 cannot be introduced into set theory because it creates a logical contradiction known as **Russell's paradox**.
  - Discovered by philosopher and logician Bertrand Russell (1872-1970) in 1901.
- **Russell's paradox**: "Let $P(x)$ be the statement... "x is a set, and $x \notin x$"; i.e., $P(x)$ is true only when x is a set which does not contain itself. For instance, $P(\{2, 3, 4\})$ is true, since the set $\{2, 3, 4\}$ is not one of the three elements 2, 3, 4 of $\{2, 3, 4\}$. On the other hand, if we let S be the set of all sets (which we would know to exist from the axiom of universal specification), then since S is itself a set, it is an element of S, and so $P(S)$ is false. Now use the axiom of universal specification to create the set

  $$\Omega := \{x : P(x) \text{ is true}\} = \{x : x \text{ is a set and } x \notin x\}$$

  i.e., the set of all sets which do not contain themselves. Now ask the question: does Ω contain itself, i.e. is $\Omega \in \Omega$? If Ω did contain itself, then by definition this means that $P(\Omega)$ is true, i.e., Ω is a set and $\Omega \notin \Omega$. On the other hand, if Ω did not contain itself, then $P(\Omega)$ would be true, and hence $\Omega \in \Omega$. Thus in either case we have both $\Omega \in \Omega$ and $\Omega \notin \Omega$, which is absurd" (Tao, 2016, pp. 46–47).
  - To clarify the last point: Is $\Omega \in \Omega$? Suppose $\Omega \in \Omega$. Then since Ω contains only sets for which $P(x)$ is true, $P(\Omega)$ must be true. But this implies, by the definition of $P(x)$, that $\Omega \notin \Omega$. Similarly, suppose $\Omega \notin \Omega$. Then since "Ω is a set and $\Omega \notin \Omega$" is a true statement, $P(\Omega)$ must be true. But this implies, since Ω contains all sets for which $P(x)$ is true, that $\Omega \in \Omega$. In either case, we have both $\Omega \in \Omega$ and $\Omega \notin \Omega$ (contradictions).

7/23:

- The main problem highlighted by Russell's paradox is that Axiom 3.8 allows for the creation of sets that are too "large," i.e., sets that contain themselves, which is somewhat silly.
  - This problem can be informally resolved by creating a hierarchy: primitive objects are below primitive sets (which only contain primitive objects), are below second-level sets (which only contain primitive objects and primitive sets), and so on and so forth. Formalizing this notion is complicated and will not be explored further here.

  - Note that in pure set theory, there are no primitive objects — only one primitive set (the empty set).

- To avoid the complications of Russell's paradox, we create a new axiom.

  Axiom 3.9 (Regularity)**.** *If A is a non-empty set, then there is at least one element x of A which is either not a set or is disjoint from A.*

- *Also known as* **axiom of foundation**.

- Axiom 3.9 asserts that "at least one of the elements of A is so low on the hierarchy of objects that it does not contain any of the other elements of A" (Tao, 2016, p. 48). It also asserts that sets may not contain themselves (see Exercise 3.2.2).

- As a less intuitive axiom, one might question whether or not Axiom 3.9 is needed. In fact, it is not necessary for the purposes of doing analysis, as all sets considered in analysis are very low on the hierarchy. However, it is necessary to perform more advanced set theory, so Tao, 2016 included it for the sake of completeness.

Exercises

7/22: 1. Show that the universal specification axiom, Axiom 3.8, if assumed to be true, would imply Axioms 3.2, 3.3, 3.4, 3.5, and 3.6. (If we assume that all natural numbers are objects, we also obtain Axiom 3.7.) Thus, this axiom, if permitted, would simplify the foundations of set theory tremendously (and can be viewed as one basis for an intuitive model of set theory known as "naive set theory"). Unfortunately, as we have seen, Axiom 3.8 is "too good to be true!"

Proof. Axiom 3.2: Let $P(x)$ be a false statement for all objects x. By Axiom 3.8, there exists a set $\{x : P(x) \text{ is true}\}$, which contains no elements. (Suppose for the sake of contradiction that $y \in \{x : P(x) \text{ is true}\}$ for some object y. Then $P(y)$ is true. But $P(y)$ is false by definition, a contradiction. Therefore, $y \notin \{x : P(x) \text{ is true}\}$ for all objects y.) Incidentally, that contradiction proof solidifies the symbolic statement of Axiom 3.2. Lastly, this set may be denoted $\emptyset$ or $\{\}$.

Axiom 3.3: Let $P(x)$ be the statement $x = a$. By Axiom 3.8, there exists a set $\{x : P(x) \text{ is true}\} = \{x : x = a\} = \{a\}$ whose element is a. For every object y, we have $y \in \{a\}$ iff $P(y)$ is true, i.e., iff $y = a$. This set may be called the **singleton set** whose element is a. Now let $P(x)$ be the statement "$x = a$ or $x = b$." By Axiom 3.8, there exists a set $\{x : P(x) \text{ is true}\} = \{x : x = a \text{ or } x = b\} = \{a, b\}$ whose elements are a and b. For every object y, we have $y \in \{a, b\}$ iff $P(y)$ is true, i.e., iff $y = a$ or $y = b$. This set may be called the **pair set** formed by a and b.

Axiom 3.4: Let A, B be sets. Let $P(x)$ be the statement "$x \in A$ or $x \in B$." By Axiom 3.8, there exists a set $\{x : P(x) \text{ is true}\} = \{x : x \in A \text{ or } x \in B\}$. This set may be called the **union** $A \cup B$ of A and B. By Axiom 3.8, $y \in A \cup B$ iff $P(y)$ is true, i.e., iff $y \in A$ or $y \in B$. Thus, $A \cup B$ is clearly a set whose elements consist of all the elements which belong to A or B or both.

Axiom 3.6: Let A be a set. Let $P(y)$ be the statement "$P(x, y)$ is true for some $x \in A$," where $P(x, y)$ is a statement pertaining to x and y such that for each $x \in A$, there is at most one y for which $P(x, y)$ is true. By Axiom 3.8, there exists a set $\{y : P(y) \text{ is true}\} = \{y : P(x, y) \text{ is true for some } x \in A\}$. By Axiom 3.8, $z \in \{y : P(x, y) \text{ is true for some } x \in A\}$ iff $P(z)$ is true, i.e., iff $P(x, z)$ is true for some $x \in A$.

Axiom 3.5: Implied by the axiom of replacement (see Exercise 3.1.11). □

7/23: 2. Use the axiom of regularity (and the singleton set axiom) to show that if A is a set, then $A \notin A$. Furthermore, show that if A and B are two sets, then either $A \notin B$ or $B \notin A$ (or both).

Proof. Suppose for the sake of contradiction that A is a set and $A \in A$. By Axioms 3.1 and 3.3, $A \in \{A\}$. Since $A \in A$ and $A \in \{A\}$, Definition 3.3 tells us that $A \in A \cap \{A\}$. Since there exists an

object x (namely A) such that $x \in A \cap \{A\}$, by Axiom 3.2 and Definition 3.1, $A \cap \{A\} \neq \emptyset$. Thus, we have A is a set and $A \cap \{A\} \neq \emptyset$. But by Axiom 3.9, as the only element of $\{A\}$, A must either not be a set or satisfy $A \cap \{A\} = \emptyset$, a contradiction. Therefore, A is not a set or $A \notin A$. Thus, if A is a set, then $A \notin A$.

Suppose for the sake of contradiction that for two sets A, B, $A \in B$ and $B \in A$. By Axioms 3.1 and 3.3, there exists a set $\{A, B\}$ whose only elements are A and B. Since $A \in B$ and $A \in \{A, B\}$, Definition 3.3, tells us that $A \in B \cap \{A, B\}$. Since there exists an object x (namely A) such that $x \in B \cap \{A, B\}$, by Axiom 3.2 and Definition 3.1, $B \cap \{A, B\} \neq \emptyset$. By a similar argument, $A \cap \{A, B\} \neq \emptyset$. Thus, we have A, B are sets, $B \cap \{A, B\} \neq \emptyset$, and $A \cap \{A, B\} \neq \emptyset$. But by Axiom 3.9, an element x of $\{A, B\}$ (namely A or B) must either not be a set or satisfy $x \cap \{A, B\} = \emptyset$, a contradiction. Therefore, $A \notin B$ or $B \notin A$. □

3. Show (assuming the other axioms of set theory) that the universal specification axiom, Axiom 3.8, is equivalent to an axiom postulating the existence of a "universal set" Ω consisting of all objects (i.e., for all objects x, we have $x \in \Omega$). In other words, if Axiom 3.8 is true, then a universal set exists, and conversely, if a universal set exists, then Axiom 3.8 is true. (This may explain why Axiom 3.8 is called the axiom of *universal* specification). Note that if a universal set Ω existed, then we would have $\Omega \in \Omega$ by Axiom 3.1, contradicting Exercise 3.2.2. Thus, the axiom of foundation specifically rules out the axiom of universal specification.

   Proof. Suppose Axiom 3.8 is true. Let $P(x)$ be a true statement for all objects x. Then there exists a set $\Omega := \{x : P(x) \text{ is true}\}$, and we have $y \in \Omega$ iff $P(y)$ is true, i.e., iff y is an object, i.e., for all objects y. Therefore, a universal set exists. Now suppose that a universal set Ω exists. By Axiom 3.5, there exists a set $\{x \in \Omega : P(x) \text{ is true}\}$ for some property $P(x)$ pertaining to x (note that this implies that $P(x)$ pertains to all x) and $y \in \{x \in \Omega : P(x) \text{ is true}\}$ iff $y \in \Omega$ and $P(y)$ is true. Since $x \in \Omega$ and $y \in \Omega$ are, by the definition of Ω, always true, we have $y \in \{x : P(x) \text{ is true}\}$ iff $P(y)$ is true. Therefore, Axiom 3.8 is true. □

3.3 Functions

- For analysis, we need not just the notion of a set but the notion of a function from one set to another.

  Definition 3.5 (Functions)**.** Let X, Y be sets, and let $P(x, y)$ be a property pertaining to an object $x \in X$ and an object $y \in Y$, such that for every $x \in X$, there is exactly one $y \in Y$ for which $P(x, y)$ is true (this is sometimes known as the **vertical line test**). Then we define the **function** $f : X \to Y$ defined by P on the **domain** X and **range** Y to be the object which, given any **input** $x \in X$, assigns an output $f(x) \in Y$, defined to be the unique object $f(x)$ for which $P(x, f(x))$ is true. Thus, for any $x \in X$ and $y \in Y$,
  $$y = f(x) \iff P(x, y) \text{ is true}$$

- *Also known as* **maps**, **transformations**, and **morphisms**.
  - Note, however, that a morphism "refers to a more general class of object, which may or may not correspond to actual functions, depending on the context" (Tao, 2016, p. 49).
- Functions obey the axiom of substitution.
  - Note that equal inputs imply equal outputs, but unequal inputs do not necessary ensure unequal outputs.
- It can be proven that this notion of equality is reflexive, symmetric, and transitive (see Exercise 3.3.1).
- We can now formally define the increment function: Let $X = \mathbb{N}$, $Y = \mathbb{N}$, and $P(x, y)$ be the property that $y = x{+}{+}$. By Axiom 2.4, for each $x \in \mathbb{N}$, there is exactly one y for which $P(x, y)$ is true. Thus, we can define the increment function $f : \mathbb{N} \to \mathbb{N}$ so that $f(x) = x{+}{+}$ for all x. While we cannot define a *decrement* function $g : \mathbb{N} \to \mathbb{N}$ ($0 \neq n{+}{+}$ for any $n \in \mathbb{N}$ by Axiom 2.3), we can define a decrement function $g : \mathbb{N} \setminus \{0\} \to \mathbb{N}$ (by Lemma 2.4).

- Informally: Note that while we cannot define a square root function $\sqrt{} : \mathbb{R} \to \mathbb{R}$, we can define the square root function $\sqrt{} : [0, +\infty) \to [0, +\infty)$.
- Functions can be defined **explicitly** or **implicitly**.
  - Explicit definitions: Specify the domain, range, and how one generates the output $f(x)$ from each input.
    - For example, the increment function f could be defined explicitly by saying that the domain and range of f are equal to $\mathbb{N}$, and $f(x) := x{+}{+}$ for all $x \in \mathbb{N}$.
  - Implicit definitions: Specify what property $P(x, y)$ links the input x with the output $f(x)$.
    - For example, the square root function $\sqrt{}$ was defined implicitly by the relation $(\sqrt{x})^2 = x$.
    - Note that implicit definitions are only valid if we know that for every input, there is only one output that obeys the implicit relation.
- Often the domain and range are not specified for the sake of brevity.
  - For example, we could refer to the increment function f as "the function $f(x) := x{+}{+}$," "the function $x \mapsto x{+}{+}$," "the function $x{+}{+}$," or the extremely abbreviated "$+{+}$."
  - Note, however, that too much abbreviation can be dangerous, omitting valuable or even necessary information.
- Note that while we now use parentheses to clarify the order of operations and enclose the arguments of functions and properties, the usages should be unambiguous from context.
  - For example, if a is a number, then $a(b+c)$ denotes $a \times (b+c)$, but if a is a function, then $a(b+c)$ denotes the output of a when the input is $b + c$.
  - Note that argument are sometimes subscripted — "a sequence of natural numbers $a_0, a_1, a_2, a_3, \ldots$ is, strictly speaking, a function from $\mathbb{N}$ to $\mathbb{N}$, but is denoted by $n \mapsto a_n$ rather than $n \mapsto a(n)$" (Tao, 2016, p. 51).
- Note that functions are not sets and sets are not functions (no $x \in f$ and $A : X \nrightarrow Y$), but we can start with a function $f : X \to Y$ and construct its **graph** $\{(x, f(x)) : x \in X\}$, which describes the function completely (see Section **??** for more).
- We now define equality for functions.

  Definition 3.6 (Equality of functions)**.** Two functions $f : X \to Y$, $g : X \to Y$ with the same domain and range are said to be equal, $f = g$, if and only if $f(x) = g(x)$ for all $x \in X$. (If $f(x)$ and $g(x)$ agree for some values of x, but not others, then we do not consider f and g to be equal.)

- Note that functions can be equal over only a certain domain.
  - For example, $x \mapsto x$ and $x \mapsto |x|$ are equal if defined only on the positive real axis, and are not equal if defined on $\mathbb{R}$.
- **Empty function**: The function $f : \emptyset \to X$.
  - We need not specify what f does to any input (since there are none), and Definition 3.6 asserts that for each set X, there is only one function from $\emptyset$ to X.
- A fundamental operation of functions is composition.

  Definition 3.7 (Composition)**.** Let $f : X \to Y$ and $g : Y \to Z$ be two functions, such that the range of f is the same set as the domain of g. We then define the **composition** $g \circ f : X \to Z$ of the two functions g and f to be the function defined explicitly by the formula

$$(g \circ f)(x) := g(f(x))$$

- It can be proven that composition obeys the axiom of substitution (see Exercise 3.3.1).
- Composition is not commutative, but it is associative.

  Lemma 3.7 (Composition is associative)**.** *Let $f : Z \to W$, $g : Y \to Z$, and $h : X \to Y$ be functions. Then $f \circ (g \circ h) = (f \circ g) \circ h$.*

  Proof. Since $g \circ h$ is a function from X to Z, $f \circ (g \circ h)$ is a function from X to W. Similarly, $f \circ g$ is a function from $Y \to W$, and hence $(f \circ g) \circ h$ is a function from $X \to W$. Thus, $f \circ (g \circ h)$ and $(f \circ g) \circ h$ have the same domain and range. In order to check that they are equal, we see from Definition 3.6 that we have to verify that $(f \circ (g \circ h))(x) = ((f \circ g) \circ h)(x)$ for all $x \in X$. But by Definition 3.7, we have

$$\begin{aligned}(f \circ (g \circ h))(x) &= f((g \circ h)(x)) \\ &= f(g(h(x))) \\ &= (f \circ g)(h(x)) \\ &= ((f \circ g) \circ h)(x)\end{aligned}$$

  as desired. □

7/24:

- We now define several special types of functions, beginning with the following.

  Definition 3.8 (One-to-one functions)**.** A function f is **one-to-one** if different elements map to different elements:

$$x \neq x' \implies f(x) \neq f(x')$$

  Equivalently, a function is one-to-one if

$$f(x) = f(x') \implies x = x'$$

- *Also known as* **injective** (function).
- Informally: Note that while the function $f : \mathbb{Z} \to \mathbb{Z}$ defined by $f(n) := n^2$ is not one-to-one, the function $g : \mathbb{N} \to \mathbb{Z}$ defined by $g(n) := n^2$ is one-to-one. Thus, being one-to-one can depend not just on the relation, but on the domain.
- **Two-to-one** (function): A function $f : X \to Y$ such that one can find distinct x and x' in the domain X such that $f(x) = f(x')$.
- Another special type is given by the following.

  Definition 3.9 (Onto functions)**.** A function f is **onto** if $f(X) = Y$, i.e., every element in Y comes from applying f to some element in X:

$$\text{For every } y \in Y, \text{ there exists } x \in X \text{ such that } f(x) = y$$

- *Also known as* **surjective** (function).
- Informally: Note that while the function $f : \mathbb{Z} \to \mathbb{Z}$ defined by $f(n) := n^2$ is not onto, if we define the set $A := \{n^2 : n \in \mathbb{Z}\}$, then the function $g : \mathbb{Z} \to A$ defined by $f(n) = n^2$ is onto. Thus, being onto can depend not just on the relation, but on the range.
- Injectivity and surjectivity are rather dual to each other (see Exercises 3.3.2, 3.3.4, and 3.3.5).

7/28:

- A third special type is given by the following.

  Definition 3.10 (Bijective functions)**.** A function $f : X \to Y$ is **bijective** if it is both one-to-one and onto:

$$\text{For every } y \in Y, \text{ there is exactly one } x \text{ such that } f(x) = y.$$

- *Also known as* **invertible** (function), **perfect matching**, **one-to-one correspondence**[3].
  - Instead of being denoted $x \mapsto f(x)$, we sometimes use $x \leftrightarrow f(x)$.
- Bijectivity of functions:
  - $f : \{0, 1, 2\} \to \{3, 4\}$ defined by $f(0) := 3$, $f(1) := 3$, and $f(2) := 4$ is not bijective (fails injectivity).
  - $f : \{0, 1\} \to \{2, 3, 4\}$ defined by $f(0) := 2$ and $f(1) := 3$ is not bijective (fails surjectivity).
  - $f : \{0, 1, 2\} \to \{3, 4, 5\}$ defined by $f(0) := 3$, $f(1) := 4$, and $f(2) := 5$ is bijective.
  - $f : \mathbb{N} \to \mathbb{N}$ defined by $f(n) := n{+}{+}$ is not bijective (fails surjectivity).
  - $f : \mathbb{N} \to \mathbb{N} \setminus \{0\}$ defined by $f(n) := n{+}{+}$ is bijective.
  - Bijectivity depends on the domain and range, not just the relation.
- **Inverse** (of a function $f : X \to Y$): The function $f^{-1} : Y \to X$ associated with the property $P(y, x)$ defined by $f(x) = y$, i.e., we let $f^{-1}(y)$ be the entity x satisfying $f(x) = y$.

Exercises

7/23: 1. Show that the definition of equality in Definition 3.6 is reflexive, symmetric, and transitive[4]. Also verify the substitution property: if $f, \tilde{f} : X \to Y$ and $g, \tilde{g} : Y \to Z$ are functions such that $f = \tilde{f}$ and $g = \tilde{g}$, then $g \circ f = \tilde{g} \circ \tilde{f}$.

Proof. Let $f : X \to Y$ be a function. For any object $f(x) \in Y$, we have $f(x) = f(x)$ by the reflexive axiom of equality (see Section A.7). Thus, we have $f(x) = f(x)$ for all $f(x) \in Y$, i.e., for all $x \in X$. Therefore, by Definition 3.6, we have $f = f$.

Let $f : X \to Y$, $g : X \to Y$ be functions, and let $f = g$. By Definition 3.6, $f(x) = g(x)$ for all $x \in X$. Since equal objects are of the same type (i.e., follow the reflexive axiom of equality), we have $g(x) = f(x)$ for all $x \in X$. Therefore, by Definition 3.6, we have $g = f$.

Let $f : X \to Y$, $g : X \to Y$, $h : X \to Y$ be functions, $f = g$, and $g = h$. By Definition 3.6, $f(x) = g(x)$ for all $x \in X$ and $g(x) = h(x)$ for all $x \in X$. Since equal objects are of the same type (i.e., follow the transitive axiom of equality), we have $f(x) = h(x)$ for all $x \in X$. Therefore, by Definition 3.6, we have $f = h$.

First, we see that both $g \circ f$ and $\tilde{g} \circ \tilde{f}$ are functions from X to Z, i.e., have the same domain and range. To show that $g \circ f = \tilde{g} \circ \tilde{f}$, Definition 3.6 tells us that it will suffice to verify that $(g \circ f)(x) = (\tilde{g} \circ \tilde{f})(x)$ for all $x \in X$. To begin, we see from Definition 3.6 that $f(x) = \tilde{f}(x)$ for all $x \in X$, and that $g(y) = \tilde{g}(y)$ for all $y \in Y$. Since $f(x) \in Y$ for all $f(x)$, we have by Definition 3.7 that

$$\begin{aligned}(g \circ f)(x) &= g(f(x)) \\ &= \tilde{g}(f(x)) \\ &= \tilde{g}(\tilde{f}(x)) \\ &= (\tilde{g} \circ \tilde{f})(x)\end{aligned}$$

as desired. □

7/24: 2. Let $f : X \to Y$ and $g : Y \to Z$ be functions. Show that if f and g are both injective, then so is $g \circ f$; similarly, show that if f and g are both surjective, then so is $g \circ f$.

[3] Not to be confused with the notion of a one-to-one function.

[4] Note that since Definition 3.6 should be an axiom (should axiomatize equality and all that that entails for functions), this part of the exercise is silly (see Tao, 2020).

Proof. To prove that $g \circ f$ is injective given the injectivity of f, g, Definition 3.8 tells us that we have to verify that $x \neq x' \implies (g \circ f)(x) \neq (g \circ f)(x')$ for any distinct elements x, x' of X. By Definition 3.8, we have that if x and x' are two distinct elements of X (such that $x \neq x'$), then $f(x) \neq f(x')$. By Definition 3.5, we know that $f(x)$ and $f(x')$ are both elements of Y. Thus, we have by Definition 3.8 that $g(f(x)) \neq g(f(x'))$. Therefore, we have by Definition 3.7 that

$$x \neq x' \implies f(x) \neq f(x') \implies g(f(x)) \neq g(f(x')) \implies (g \circ f)(x) \neq (g \circ f)(x')$$

as desired.

By Definition 3.7, we have $g \circ f : X \to Z$. Thus, to prove that $g \circ f$ is surjective given the surjectivity of f, g, Definition 3.9 tells us that we have to verify that for every $z \in Z$, there exists $x \in X$ such that $(g \circ f)(x) = z$. Let z be any element of Z. Then by Definition 3.9 and the surjectivity of g, we have $g(y) = z$ for some $y \in Y$. Similarly, we have $f(x) = y$ for that y and for some $x \in X$. Substituting, we have $g(f(x)) = z$ for some $x \in X$. Since $g(f(x)) = (g \circ f)(x)$ by Definition 3.7, we have that for any $z \in Z$, there exists $x \in X$ such that $(g \circ f)(x) = z$. □

7/28: 3. When is the empty function injective? Surjective? Bijective?

Proof. The empty function is always injective: The statement "all distinct inputs map to distinct outputs" is vacuously true, since there exist no distinct inputs, or any inputs to speak of. The empty function is surjective iff its range is the empty set: In this case, the statement "for every $y \in \emptyset$, there exists $x \in \emptyset$ such that $f(x) = y$" is similarly vacuously true. Since a function is bijective iff it is both injective and surjective, we must take both take a sort of union of the two constraints above: We can conclude that the empty function is bijective iff "always and its range is the empty set" is true, i.e., iff its range is the empty set. □

7/24: 4. In this section, we give some cancellation laws for composition. Let $f : X \to Y$, $\tilde{f} : X \to Y$, $g : Y \to Z$, and $\tilde{g} : Y \to Z$ be functions. Show that if $g \circ f = g \circ \tilde{f}$ and g is injective, then $f = \tilde{f}$. Is the same statement true if g is not injective? Show that if $g \circ f = \tilde{g} \circ f$ and f is surjective, then $g = \tilde{g}$. Is the same statement true if f is not surjective?

Proof. To prove that $f = \tilde{f}$, Definition 3.6 tells us that it will suffice to show that $f(x) = \tilde{f}(x)$ for all $x \in X$. By Definition 3.6, $g \circ f = g \circ \tilde{f}$ implies $(g \circ f)(x) = (g \circ \tilde{f})(x)$ for all $x \in X$. Then by Definition 3.7, $g(f(x)) = g(\tilde{f}(x))$ for all $x \in X$. Now by Definition 3.8 and the fact that g is injective, we know that $g(y) = g(y') \implies y = y'$. Therefore, since $g(f(x)) = g(\tilde{f}(x))$ for all $x \in X$, we have $f(x) = \tilde{f}(x)$ for all $x \in X$. Note that f is not necessarily equal to $\tilde{f}$ if we drop the condition that g is injective (if g is not one-to-one, then we have $g(y) = g(y')$ for some elements y and y' of Y such that $y \neq y'$. Thus, if $f(x) \neq \tilde{f}(x)$, we may still have $g(f(x)) = g(\tilde{f}(x))$).

Suppose for the sake of contradiction that $g \neq \tilde{g}$. Then by Definition 3.6, either g and $\tilde{g}$ have a different domain or range, or $g(y) \neq \tilde{g}(y)$ for some $y \in Y$. Since g and $\tilde{g}$ have the same domain and range, we must have $g(y) \neq \tilde{g}(y)$ for some $y \in Y$. Now since f is surjective, by Definition 3.9, for every $y \in Y$, there is some $x \in X$ such that $y = f(x)$. Thus, $g(f(x)) \neq \tilde{g}(f(x))$ for some $x \in X$. Consequently, by Definition 3.7, $(g \circ f)(x) \neq (\tilde{g} \circ f)(x)$ for some $x \in X$. But this contradicts Definition 3.6, which, since $g \circ f = \tilde{g} \circ f$, implies that $(g \circ f)(x) = (\tilde{g} \circ f)(x)$ for all $x \in X$. Note that g is not necessarily equal to $\tilde{g}$ if we drop the condition that f is surjective (if f is not surjective, then, as we've only proven the assertion for all $y = f(x)$, we could have some $y \in Y$ not associated with an $f(x)$ for which $g(y) \neq \tilde{g}(y)$). □

7/28: 5. Let $f : X \to Y$ and $g : Y \to Z$ be functions. Show that if $g \circ f$ is injective, then f must be injective. Is it true that g must also be injective? Show that if $g \circ f$ is surjective, then g must be surjective. Is it true that f must also be surjective?

Proof. By Definition 3.8, to prove that f is injective, it will suffice to show that for all $x \neq x'$, $f(x) \neq f(x')$. Suppose for the sake of contradiction that for some $x \neq x'$, $f(x) = f(x')$. Then by Definition 3.5, which guarantees that identical inputs map to a single output, we have $g(f(x)) = g(f(x'))$. Thus, by Definition 3.7, we have $(g \circ f)(x) = (g \circ f)(x')$. But by Definition 3.8, this means that $g \circ f$ is not injective, a contradiction. Therefore, for all $x \neq x'$, $f(x) \neq f(x')$. Note that g need not be injective to prove the above, as this condition was not needed for the above proof to be valid.

By Definition 3.9, to prove that g is surjective, we must verify that for every $z \in Z$, there exists $y \in Y$ such that $g(y) = z$. Let z be any element of Z. Since we know that $g \circ f$ is surjective, Definition 3.9 tells us that there exists an element x of X such that $(g \circ f)(x) = z$. By Definition 3.7, we have $g(f(x)) = z$. Now by Definition 3.5, we know that $f(x) = y$ for some element $y \in Y$. Thus, $g(y) = z$ for some $y \in Y$. Therefore, we have proven that for any (i.e., for every) element $z \in Z$, there exists $y \in Y$ such that $g(y) = z$. Note that f need not be surjective to prove the above, as this condition was not needed for the above proof to be valid. □

6. Let $f : X \to Y$ be a bijective function, and let $f^{-1} : Y \to X$ be its inverse. Verify the cancellation laws $f^{-1}(f(x)) = x$ for all $x \in X$ and $f(f^{-1}(y)) = y$ for all $y \in Y$. Conclude that f^{-1} is also invertible, and has f as its inverse (thus $(f^{-1})^{-1} = f$).

Proof. Let y be any element of Y. We know that f^{-1} maps y to the element $x \in X$ satisfying $f(x) = y$ (the uniqueness of x being guaranteed by Definition 3.10). Thus, since $y = f(x)$, f^{-1} maps $f(x)$ to x, i.e., $f^{-1}(f(x)) = x$. A similar argument can treat the other cancellation law.

To prove that f^{-1} is invertible, we must verify that for every $x \in X$, there is exactly one y such that $f^{-1}(y) = x$. Since we know that for every $x \in X$, $f^{-1}(f(x)) = x$, and $f(x) = y$ for some $y \in Y$ (Definition 3.5), we know that for every $x \in X$, $f^{-1}(y) = x$ for some $y \in Y$. Now suppose for the sake of contradiction that there exist some $y \neq y'$ such that $f^{-1}(y) = x = f^{-1}(y')$. Then $f(x) = y$ and $f(x) = y'$, which contradicts Definition 3.5. Therefore, for every $x \in X$, $f^{-1}(y) = x$ for some and for not more than one $y \in Y$, i.e., for exactly one $y \in Y$.

The inverse of f^{-1} is, by definition, the function $(f^{-1})^{-1} : X \to Y$ associated with the property $P(x, y)$ defined by $f^{-1}(y) = x$. Clearly $(f^{-1})^{-1}$ has the same domain and range as f. All that's left is to show that $f(x) = (f^{-1})^{-1}(x)$. $f(x)$ is the element $y \in Y$ such that $f(x) = y$, and $(f^{-1})^{-1}(x)$ is the element $y \in Y$ such that $f^{-1}(y) = x$, i.e., $f(x) = y$. Since the functions are associated with the same property, they must be equal. □

7. Let $f : X \to Y$ and $g : Y \to Z$ be functions. Show that if f and g are bijective, then so is $g \circ f$, and we have $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.

Proof. By Definition 3.10, f and g are both both injective and surjective. Thus, by consecutive applications of Exercise 3.3.2, $g \circ f$ is both injective and surjective. Thus, by Definition 3.10 again, $g \circ f$ is bijective.

To prove that $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$, Definition 3.6 tells us that it will suffice to show that $(g \circ f)^{-1}$ and $f^{-1} \circ g^{-1}$ have the same domain and range, and $(g \circ f)^{-1}(z) = (f^{-1} \circ g^{-1})(z)$ for all $z \in Z$. Since $f : X \to Y$ and $g : Y \to Z$, by Definition 3.7, we have $g \circ f : X \to Z$. Thus, the inverse $(g \circ f)^{-1} : Z \to X$. Similarly, since $g^{-1} : Z \to Y$ and $f^{-1} : Y \to X$, by Definition 3.7, $f^{-1} \circ g^{-1} : Z \to X$. Thus, $(g \circ f)^{-1}$ and $f^{-1} \circ g^{-1}$ have the same domain and range. As to the other part of the question, let $(f^{-1} \circ g^{-1})(z) = x$. Then

$$\begin{aligned}
(f^{-1} \circ g^{-1})(z) = x &\implies f^{-1}(g^{-1}(z)) = x && \text{Definition 3.7} \\
&\implies g^{-1}(z) = f(x) && \text{Definition of inverse} \\
&\implies z = g(f(x)) && \text{Definition of inverse} \\
&\implies z = (g \circ f)(x) && \text{Definition 3.7} \\
&\implies (g \circ f)^{-1}(z) = x && \text{Definition of inverse}
\end{aligned}$$

so by transitivity we have $(g \circ f)^{-1}(z) = (f^{-1} \circ g^{-1})(z)$, as desired. □

8. If X is a subset of Y, let $\iota_{X\to Y} : X \to Y$ be the **inclusion map** (from X to Y), defined by mapping $x \mapsto x$ for all $x \in X$, i.e., $\iota_{X\to Y}(x) := x$ for all $x \in X$. The map $\iota_{X\to X}$ is in particular called the **identity map** (on X).

   (a) Show that if $X \subseteq Y \subseteq Z$, then $\iota_{Y\to Z} \circ \iota_{X\to Y} = \iota_{X\to Z}$.

   Proof. By Definition 3.7, $\iota_{Y\to Z} \circ \iota_{X\to Y} : X \to Z$, so $\iota_{Y\to Z} \circ \iota_{X\to Y}$ and $\iota_{X\to Z}$ have the same domain and range. Now we must check that $\iota_{Y\to Z} \circ \iota_{X\to Y}$ and $\iota_{X\to Z}$ match the same inputs to the same outputs. Since $X \subseteq Y \subseteq Z$, $X \subseteq Z$ (Proposition 3.1), so $\iota_{X\to Z}(x) = x$ for all $x \in X$. On the other hand, $X \subseteq Y$ guarantees that $\iota_{X\to Y}(x) = x$ for all $x \in X$ and $Y \subseteq Z$ guarantees that $\iota_{Y\to Z}(y) = y$ for all $y \in Y$. Since $x \in Y$ for all $x \in X$ (Definition 3.2),

$$\begin{aligned} \iota_{Y\to Z} \circ \iota_{X\to Y}(x) &= \iota_{Y\to Z}(\iota_{X\to Y}(x)) && \text{Definition 3.7} \\ &= \iota_{Y\to Z}(x) \\ &= x \end{aligned}$$

   for all $x \in X$. Therefore, $\iota_{Y\to Z} \circ \iota_{X\to Y} = \iota_{X\to Z}$. □

   (b) Show that if $f : A \to B$ is any function, then $f = f \circ \iota_{A\to A} = \iota_{B\to B} \circ f$.

   Proof. We first prove that $f = f \circ \iota_{A\to A}$, and next prove that $f = \iota_{B\to B} \circ f$. Transitivity will guarantee that $f = f \circ \iota_{A\to A} = \iota_{B\to B} \circ f$.
   To prove that $f = f \circ \iota_{A\to A}$, Definition 3.6 tells us that it will suffice to show that f and $f \circ \iota_{A\to A}$ have the same domain and range, and that $f(a) = (f \circ \iota_{A\to A})(a)$ for all $a \in A$. By Definition 3.7, $f \circ \iota_{A\to A} : A \to B$, so f and $f \circ \iota_{A\to A}$ have the same domain and range. Now since $A \subseteq A$, $\iota_{A\to A}(a) = a$ for all $a \in A$. Thus, by Definition 3.7, $(f \circ \iota_{A\to A})(a) = f(\iota_{A\to A}(a)) = f(a)$.
   To prove that $f = \iota_{B\to B} \circ f$, Definition 3.6 tells us that it will suffice to show that f and $\iota_{B\to B} \circ f$ have the same domain and range, and that $f(a) = (\iota_{B\to B} \circ f)(a)$ for all $a \in A$. By Definition 3.7, $\iota_{B\to B} \circ f : A \to B$, so f and $\iota_{B\to B} \circ f$ have the same domain and range. Now since $B \subseteq B$, $\iota_{B\to B}(b) = b$ for all $b \in B$. Since every $f(a) \in B$, by Definition 3.7, $(\iota_{B\to B} \circ f)(a) = \iota_{B\to B}(f(a)) = f(a)$. □

   (c) Show that if $f : A \to B$ is a bijective function, then $f \circ f^{-1} = \iota_{B\to B}$, and $f^{-1} \circ f = \iota_{A\to A}$.

   Proof. By Exercise 3.3.6 (and Definition 3.7), $(f \circ f^{-1})(b) = b$ for all $b \in B$. By definition, $\iota_{B\to B}(b) = b$ for all $b \in B$. Therefore, $f \circ f^{-1} = \iota_{B\to B}$. A similar argument holds in the other case. □

   (d) Show that if X and Y are disjoint sets, and $f : X \to Z$ and $g : Y \to Z$ are functions, then there is a unique function $h : X \cup Y \to Z$ such that $h \circ \iota_{X\to X\cup Y} = f$ and $h \circ \iota_{Y\to X\cup Y} = g$.

   Proof. Let $h : X \cup Y \to Z$ be defined by the following (note that it is the fact that X and Y are disjoint that allows the following definition to make sense, i.e., not be contradictory for some inputs).

$$h(a) := \begin{cases} f(a) & a \in X \\ g(a) & a \in Y \end{cases}$$

   To prove that $h \circ \iota_{X\to X\cup Y} = f$, Definition 3.6 tells us that it will suffice to show that $h \circ \iota_{X\to X\cup Y}$ and f have the same domain and range, and that $(h \circ \iota_{X\to X\cup Y})(a) = f(a)$ for all $a \in X$. By Definition 3.7, $h \circ \iota_{X\to X\cup Y} : X \to Z$, so $h \circ \iota_{X\to X\cup Y}$ and f have the same domain and range. Now since $X \subseteq X \cup Y$ (Exercise 3.1.7), $\iota_{X\to X\cup Y}(a) = a$ for all $a \in X$. Thus, by Definition 3.7, for any $a \in X$, $(h \circ \iota_{X\to X\cup Y})(a) = h(\iota_{X\to X\cup Y}(a)) = h(a) = f(a)$. A similar argument holds for the other case[5]. □

[5] How do I verify uniqueness, or do I not need to?

3.4 Images and Inverse Images

7/29:

- We now discuss what happens when we apply a function to a set, as opposed to individual, non-set objects.

  Definition 3.11 (Images of sets)**.** If $f : X \to Y$ is a function from X to Y, and S is a set in X (or a subset of X), we define $f(S)$ to be the set

  $$f(S) := \{f(x) : x \in S\}$$

  This set is a subset of Y, and is sometimes called the **image** (of S under the map f). We sometimes call $f(S)$ the **forward image** (of S) to distinguish it from the concept of the **inverse image** $f^{-1}(S)$ (of S), which is defined below.

  Note that, symbolically,

  $$x \in S \implies f(x) \in f(S)$$

  and

  $$y \in f(S) \iff y = f(x) \text{ for some } x \in S$$

- Note that $f(S)$ is well-defined, as Definition 3.11 is based in Axiom 3.6. Also note that Definition 3.11 could be defined in terms of Axiom 3.5.

- As alluded to above, we now define inverse images.

  Definition 3.12 (Inverse images)**.** If U is a subset of Y, we define the set $f^{-1}(U)$ to be the set

  $$f^{-1}(U) := \{x \in X : f(x) \in U\}$$

  In other words, $f^{-1}(U)$ consists of all the elements of X which map into U:

  $$f(x) \in U \iff x \in f^{-1}(U)$$

  We call $f^{-1}(U)$ the **inverse image** (of U).

- Note that if f is bijective, then we have defined f^{-1} in two different ways. However, the definitions are equivalent (see Exercise 3.4.1).

- Since we wish to treat functions as objects so that we can create sets of functions, we axiomatize this notion.

  Axiom 3.10 (Power set axiom)**.** *Let X and Y be sets. Then there exists a set, denoted Y^X, which consists of all the functions from X to Y, thus*

  $$f \in Y^X \iff f \text{ is a function with domain } X \text{ and range } Y$$

- "Let $X = \{4, 7\}$ and $Y = \{0, 1\}$. Then the set Y^X consists of four functions: the function that maps $4 \mapsto 0$ and $7 \mapsto 0$; the function that maps $4 \mapsto 0$ and $7 \mapsto 1$; the function that maps $4 \mapsto 1$ and $7 \mapsto 0$; and the function that maps $4 \mapsto 1$ and $7 \mapsto 1$" (Tao, 2016, p. 58).

- Note that we use the notation Y^X because if Y has n elements and X has m elements, then Y^X has n^m elements (see Proposition **??**).

- See Exercise 3.4.6 for a consequence of Axiom 3.10.

- **Power set** (of X): The set $2^X = \{Y : Y \text{ is a subset of } X\}$.

- Note that we use the notation 2^X because if X has m elements, then 2^X has 2^m elements (see Chapter **??**).

- For the sake of completeness, we enhance Axiom 3.4 with the following.

Axiom 3.11 (Union)**.** *Let A be a set, all of whose elements are themselves sets. Then there exists a set $\bigcup A$ whose elements are precisely those objects which are elements of the elements of A. Thus, for all objects x,*

$$x \in \bigcup A \iff x \in S \text{ for some } S \in A$$

- Note that Axioms 3.11 and 3.3 imply Axiom 3.4 (see Exercise 3.4.8).
- An important consequence of Axiom 3.11 is that "if one has some set I, and for every element $\alpha \in I$ we have some set A_α, then we can form the union set $\bigcup_{\alpha \in I} A_\alpha$ by defining

$$\bigcup_{\alpha \in I} A_\alpha := \bigcup \{A_\alpha : \alpha \in I\} \tag{3.1}$$

  which is a set thanks to the axiom of replacement and the axiom of union" (Tao, 2016, p. 59).
- Note that for any object y,

$$y \in \bigcup_{\alpha \in I} A_\alpha \iff y \in A_\alpha \text{ for some } \alpha \in I \tag{3.2}$$

- To explain the above, consider the following example.
  - If $I = \{1, 2, 3\}$, $A_1 := \{2, 3\}$, $A_2 := \{3, 4\}$, and $A_3 := \{4, 5\}$, then $\bigcup_{\alpha \in \{1,2,3\}} A_\alpha = \{2, 3, 4, 5\}$.
- There is some notation associated with this paradigm, defined as follows.
- **Index set**: The set I.
- **Labels**: The elements α of the index set I.
- **Family of sets**: The group of sets A_α for all $\alpha \in I$.
  - Note that the family of sets is **indexed** by the labels $\alpha \in I$.
- Note that if I were empty, then $\bigcup_{\alpha \in I} A_\alpha$ would also be empty.
- "Given any non-empty set I, and given an assignment of a set A_α to each $\alpha \in I$, we can define the intersection $\bigcap_{\alpha \in I} A_\alpha$ by first choosing some element β of I (which we can do since I is nonempty), and setting

$$\bigcap_{\alpha \in I} A_\alpha := \{x \in A_\beta : x \in A_\alpha \text{ for all } \alpha \in I\} \tag{3.3}$$

  which is a set by the axiom of specification" (Tao, 2016, p. 60).
- Note that for any object y,

$$y \in \bigcap_{\alpha \in I} A_\alpha \iff y \in A_\alpha \text{ for all } \alpha \in I \tag{3.4}$$

- Note that this definition does not depend on the choice of β (see Exercise 3.4.9).
- Note that Axioms 3.1-3.11 (excluding Axiom 3.8) are known as the **Zermelo-Fraenkel axioms of set theory**, after Ernest Zermello (1871-1953) and Abraham Fraenkel (1891-1965).
  - Note that "these axioms are formulated slightly differently in other texts, but all the formulations can be shown to be equivalent to each other" (Tao, 2016, p. 60).
  - Also note that there is one additional axiom (the **axiom of choice** — see Section **??**), giving rise to the **Zermelo-Fraenkel-Choice (ZFC) axioms of set theory**, but we do not need this axiom at this moment.

Exercises

1. Let $f : X \to Y$ be a bijective function, and let $f^{-1} : Y \to X$ be its inverse. Let V be any subset of Y. Prove that the forward image of V under f^{-1} is the same set as the inverse image of V under f; thus, the fact that both sets are denoted by $f^{-1}(V)$ will not lead to any inconsistency.

   Proof. By Definition 3.11, the forward image of V under f^{-1} is the set $\{f^{-1}(x) : x \in V\}$. By Definition 3.12, the inverse image of V under f is the set $\{x \in X : f(x) \in V\}$. Now to prove that these sets are equal, by Definition 3.1, we must verify that every element y of $\{f^{-1}(x) : x \in V\}$ is an element of $\{x \in X : f(x) \in V\}$ and vice versa. Suppose first that y is an arbitrary element of $\{f^{-1}(x) : x \in V\}$. By Definition 3.11, this implies that $y = f^{-1}(x)$ for some $x \in V$. Thus, by the definition of the inverse (and the fact that f is bijective), $f(y) = x$ for some $x \in V$, or, more simply, $f(y) \in V$. But by Definition 3.12, $f(y) \in V \implies y \in \{x \in X : f(x) \in V\}$. Using the above implications in reverse will suffice to prove that $y \in \{x \in X : f(x) \in V\} \implies y \in \{f^{-1}(x) : x \in V\}$. □

7/30:

2. Let $f : X \to Y$ be a function from one set X to another set Y, let S be a subset of X, and let U be a subset of Y. What, in general, can one say about $f^{-1}(f(S))$ and S? What about $f(f^{-1}(U))$ and U?

   Proof. It is hard to generalize much from the rigid definitions, but we will give those and then prove one property. So first off,

   $$f^{-1}(f(S)) = \{x \in X : f(x) \in \{f(x') : x' \in S\}\}$$
   $$f(f^{-1}(U)) = \{f(x) : x \in \{x' \in X : f(x') \in U\}\}$$

   Now, we can actually prove that $S \subseteq f^{-1}(f(S))$ and $f(f^{-1}(U)) \subseteq U$, which we will do as follows.

   By Definition 3.2, to prove that $S \subseteq f^{-1}(f(S))$, we must verify that every element $x \in S$ is an element of $f^{-1}(f(S))$. Suppose x is any element of S. To prove that $x \in \{x' \in X : f(x') \in \{f(x'') : x'' \in S\}\}$, Axiom 3.5 tells us that it will suffice to show that $x \in X$ and "$f(x) \in \{f(x') : x' \in S\}$" is a true statement. Since $S \subseteq X$ and $x \in S$, we have by Definition 3.2 that $x \in X$. On the other hand, let $y = f(x)$. Then since $x \in S$, by Axiom 3.6, $y \in \{y' : y' = f(x') \text{ for some } x' \in S\} \implies f(x) \in \{f(x') : x' \in S\}$.

   By Definition 3.2, to prove that $f(f^{-1}(U)) \subseteq U$, we must verify that every element $y \in f(f^{-1}(U))$ is an element of U. Suppose y is any element of $f(f^{-1}(U))$. Then by the rigid definition of $f(f^{-1}(U))$, above, and Definition 3.1, we have $y \in \{f(x) : x \in \{x' \in X : f(x') \in U\}\}$. By Axiom 3.6, this means that $y = f(x)$ for some $x \in \{x' \in X : f(x') \in U\}$. But by Axiom 3.5, $x \in \{x' \in X : f(x') \in U\} \implies (x \in X \text{ and } f(x) \in U)$. Thus, we know that $y = f(x)$ for some $f(x) \in U$, so we have $y \in U$. □

3. Let A, B be two subsets of a set X, and let $f : X \to Y$ be a function. Show that $f(A \cap B) \subseteq f(A) \cap f(B)$, that $f(A) \setminus f(B) \subseteq f(A \setminus B)$, and that $f(A \cup B) = f(A) \cup f(B)$. For the first two statements, is it true that the $\subseteq$ relation can be improved to $=$?

   Proof. Because of the number of statements to prove and the logical simplicity of the implications involved, we will shorthand these proofs.

   We begin by proving that $f(A \cap B) \subseteq f(A) \cap f(B)$, for which it will suffice (Definition 3.2) to show that every element y of $f(A \cap B)$ is an element of $f(A) \cap f(B)$. To this end, let y be any element of $f(A \cap B)$. Since $A \cap B \subseteq A$ (Exercise 3.1.7) and $A \subseteq X$, $A \cap B \subseteq X$ (Proposition 3.1). Thus,

   $$\begin{aligned} y \in f(A \cap B) &\implies y = f(x) \text{ for some } x \in A \cap B && \text{Definition 3.11} \\ &\implies (y = f(x) \text{ for some } x \in A \text{ and } y = f(x) \text{ for some } x \in B) && \text{Definition 3.3} \\ &\implies (y \in f(A) \text{ and } y \in f(B)) && \text{Definition 3.11} \\ &\implies y \in f(A) \cap f(B) && \text{Definition 3.3} \end{aligned}$$

To prove that $f(A)\setminus f(B) \subseteq f(A\setminus B)$, we must verify (Definition 3.2) that every element y of $f(A)\setminus f(B)$ is an element of $f(A\setminus B)$, which can be done as follows. Note that $A\setminus B \subseteq X$ since $A\setminus B \subseteq A$ (Definition 3.4 and Axiom 3.5) and $A \subseteq X$ (Proposition 3.1).

$$\begin{aligned} y \in f(A) \setminus f(B) &\implies y \in \{z \in f(A) : z \notin f(B)\} && \text{Definition 3.4} \\ &\implies (y \in f(A) \text{ and } y \notin f(B)) && \text{Axiom 3.5} \\ &\implies (y = f(x) \text{ for some } x \in A \text{ and } y = f(x) \text{ for no } x \in B) && \text{Definition 3.11} \\ &\implies y = f(x) \text{ for some } x \in A \setminus B && \text{Definition 3.4} \\ &\implies y \in f(A \setminus B) && \text{Definition 3.11} \end{aligned}$$

To prove that $f(A \cup B) = f(A) \cup f(B)$, Definition 3.1 tells us that it will suffice to show that $y \in f(A \cup B) \iff y \in f(A) \cup f(B)$, which can be done as follows.

$$\begin{aligned} y \in f(A \cup B) &\iff y = f(x) \text{ for some } x \in A \cup B && \text{Definition 3.11} \\ &\iff (y = f(x) \text{ for some } x \in A \text{ or } y = f(x) \text{ for some } x \in B) && \text{Axiom 3.4} \\ &\iff (y \in f(A) \text{ or } y \in f(B)) && \text{Definition 3.11} \\ &\iff y \in f(A) \cup f(B) && \text{Axiom 3.4} \end{aligned}$$

Because of the transitivity of logically equivalent statements (we may induct the result of Exercise A.1.5), we have shown what was desired.

As to the other part of the question, it is *not* true that the $\subseteq$ relation can be improved to $=$ for the first or the second statement, which we may verify thorugh two counterexamples, as follows. For the first statement, let $X = \{1, 2\}$, $Y = \{3, 4\}$, $A = \{1\}$, $B = \{2\}$, and $f : \{1, 2\} \to \{3, 4\}$ be defined by $f(x) := 3$. Then we have $f(A) \cap f(B) = \{3\}$, but $f(A \cap B) = \{\}$. Thus, $3 \in f(A) \cap f(B)$, but $3 \notin f(A \cap B)$. Note that the break down in the would-be reverse proof of $f(A \cap B) \subseteq f(A) \cap f(B)$ comes from the fact that although $y \in f(A) \cap f(B) \implies y = f(x)$ for some $x \in A$ and $y = f(x')$ for some $x' \in B$, we cannot guarantee that $x = x'$, i.e., that $x \in A \cap B$. As to the second statement, we consider the same function, except that we let $A = \{1, 2\}$. Thus, we have $f(A \setminus B) = \{3\}$, but $f(A) \setminus f(B) = \{\}$, implying that $3 \in f(A \setminus B)$ but $3 \notin f(A) \setminus f(B)$. Note that the breakdown in the would-be reverse proof of $f(A) \setminus f(B) \subseteq f(A \setminus B)$ comes from the fact that although $y \in f(A \setminus B) \implies y = f(x)$ for some $x \in A$, we cannot guarantee that there exists no $x' \in B$ such that $y = f(x')$. □

4. Let $f : X \to Y$ be a function from one set X to another set Y, and let U, V be subsets of Y. Show that $f^{-1}(U \cup V) = f^{-1}(U) \cup f^{-1}(V)$, that $f^{-1}(U \cap V) = f^{-1}(U) \cap f^{-1}(V)$, and that $f^{-1}(U \setminus V) = f^{-1}(U) \setminus f^{-1}(V)$.

Proof. For the same reasons as in Exercise 3.4.3, we shorthand these proofs.

To prove that $f^{-1}(U \cup V) = f^{-1}(U) \cup f^{-1}(V)$, Definition 3.1 tells us that it will suffice to show that $x \in f^{-1}(U \cup V) \iff x \in f^{-1}(U) \cup f^{-1}(V)$, which can be done as follows.

$$\begin{aligned} x \in f^{-1}(U \cup V) &\iff f(x) \in U \cup V && \text{Definition 3.12} \\ &\iff (f(x) \in U \text{ or } f(x) \in V) && \text{Axiom 3.4} \\ &\iff (x \in f^{-1}(U) \text{ or } x \in f^{-1}(V)) && \text{Definition 3.12} \\ &\iff x \in f^{-1}(U) \cup f^{-1}(V) && \text{Axiom 3.4} \end{aligned}$$

To prove that $f^{-1}(U \cap V) = f^{-1}(U) \cap f^{-1}(V)$, Definition 3.1 tells us that it will suffice to show that $x \in f^{-1}(U \cap V) \iff x \in f^{-1}(U) \cap f^{-1}(V)$, which can be done as follows.

$$\begin{aligned} x \in f^{-1}(U \cap V) &\iff f(x) \in U \cap V && \text{Definition 3.12} \\ &\iff (f(x) \in U \text{ and } f(x) \in V) && \text{Definition 3.3} \\ &\iff (x \in f^{-1}(U) \text{ and } x \in f^{-1}(V)) && \text{Definition 3.12} \\ &\iff x \in f^{-1}(U) \cap f^{-1}(V) && \text{Definition 3.3} \end{aligned}$$

To prove that $f^{-1}(U \setminus V) = f^{-1}(U) \setminus f^{-1}(V)$, Definition 3.1 tells us that it will suffice to show that $x \in f^{-1}(U \setminus V) \iff x \in f^{-1}(U) \setminus f^{-1}(V)$, which can be done as follows.

$$\begin{aligned}
x \in f^{-1}(U \setminus V) &\iff f(x) \in U \setminus V && \text{Definition 3.12} \\
&\iff f(x) \in \{y \in U : y \notin V\} && \text{Definition 3.4} \\
&\iff (f(x) \in U \text{ and } f(x) \notin V) && \text{Axiom 3.5} \\
&\iff (x \in f^{-1}(U) \text{ and } x \notin f^{-1}(V)) && \text{Definition 3.12} \\
&\iff x \in \{f^{-1}(U) : x \notin f^{-1}(V)\} && \text{Axiom 3.5} \\
&\iff x \in f^{-1}(U) \setminus f^{-1}(V) && \text{Definition 3.12}
\end{aligned}$$

□

5. Let $f : X \to Y$ be a function from one set X to another set Y. Show that $f(f^{-1}(S)) = S$ for every $S \subseteq Y$ if and only if f is surjective. Show that $f^{-1}(f(S)) = S$ for every $S \subseteq X$ if and only if f is injective.

Proof. To show that $f(f^{-1}(S)) = S$ for every $S \subseteq Y$ if and only if f is surjective, we must verify that the implication "if $f(f^{-1}(S)) = S$ for every $S \subseteq Y$, then f is surjective" and its converse are both true.

Suppose first that $f(f^{-1}(S)) = S$ for every $S \subseteq Y$. To show that f is surjective, Definition 3.9 tells us that it will suffice to show that for every $y \in Y$, there exists $x \in X$ such that $f(x) = y$. Let y be any element of Y. Then by Axiom 3.3, there exists a set $\{y\}$. Since $y \in Y$ and y is the only element of $\{y\}$, by Definition 3.2, $\{y\} \subseteq Y$. Thus, $f(f^{-1}(\{y\})) = \{y\}$. Since $\{y\}$ is a nonempty set, we know that $f^{-1}(\{y\})$ is a nonempty set (if $f^{-1}(\{y\})$ were empty, then $f(f^{-1}(\{y\})) = \{y\}$ would be empty, a contradiction). Thus, there exists some $x \in f^{-1}(\{y\})$. Since Definition 3.12 implies that $x \in \{x' \in X : f(x') \in \{y\}\}$, meaning by Axiom 3.5 that $x \in X$ and $f(x) \in \{y\}$, and, thus, by Axiom 3.3 that $f(x) = y$, we have that there exists $x \in X$ such that $f(x) = y$, as desired.

Now suppose that f is surjective. To show that $f(f^{-1}(S)) = S$ for every $S \subseteq Y$, we must verify that for any subset S of Y, every element y of $f(f^{-1}(S))$ is an element of S and vice versa (Definition 3.1). Let S be an arbitrary subset of Y. By Exercise 3.4.2, $f(f^{-1}(S)) \subseteq S$. On the other hand, suppose that y is an arbitrary element of S. Since f is surjective, we have by Definition 3.9 that $f(x) = y$ for some $x \in X$. Now consider this x: since $y \in S$, we know that $f(x) \in S$. Thus, by Definition 3.12, $x \in f^{-1}(S)$. Therefore, we have $y = f(x)$ for some $x \in f^{-1}(S)$, which by Definition 3.11 and the fact that $f^{-1}(S) \subseteq X$ (as implied by Definition 3.12 and Axiom 3.5) implies that $y \in f(f^{-1}(S))$.

To show that $f^{-1}(f(S)) = S$ for every $S \subseteq X$ if and only if f is injective, we must verify that the implication "if $f^{-1}(f(S)) = S$ for every $S \subseteq X$, then f is injective" and its converse are both true.

Suppose first that $f^{-1}(f(S)) = S$ for every $S \subseteq X$. To show that f is injective, Definition 3.8 tells us that it will suffice to show that if $x \neq x'$ for two elements x, x' of X, then $f(x) \neq f(x')$. Let x, x' be any two elements of X such that $x \neq x'$. By Axiom 3.3, the sets $\{x\}$ and $\{x'\}$ exist. Note that since $x \in X$ and x is the only element of $\{x\}$, by Definition 3.2, $\{x\} \subseteq X$, and, similarly, $\{x'\} \subseteq X$. Since there exists an element of $\{x\}$ (namely x) that is not an element of $\{x'\}$, by Definition 3.1, $\{x\} \neq \{x'\}$. Thus, since $f^{-1}(f(\{x\})) = \{x\}$ and $f^{-1}(f(\{x'\})) = \{x'\}$, we have $f^{-1}(f(\{x\})) \neq f^{-1}(f(\{x'\}))$. This implies that $f(\{x\}) \neq f(\{x'\})$ (if $f(\{x\})$ were equal to $f(\{x'\})$, then $f^{-1}(f(\{x\}))$ would be equal to $f^{-1}(f(\{x'\}))$, a contradiction). Since $f(\{x\}) = \{f(x'') : x'' \in \{x\}\} = \{f(x)\}$, and, similarly, $f(\{x'\}) = \{f(x')\}$, we have $\{f(x)\} \neq \{f(x')\}$. Therefore, $f(x) \neq f(x')$ (if $f(x) = f(x')$, then, as singleton sets (Axiom 3.3), every element of $\{f(x)\}$ would be an element of $\{f(x')\}$, and vice versa, proving their equality by Definition 3.1, a contradiction).

Now suppose that f is injective. To show that $f^{-1}(f(S)) = S$ for every $S \subseteq X$, we must verify that for any subset S of X, every element x of $f^{-1}(f(S))$ is an element of S and vice versa (Definition 3.1). Let S be an arbitrary subset of X, and suppose that x is an arbitrary element of $f^{-1}(f(S))$. First off, $f(S) \subseteq Y$. (We must show that every $y \in f(S)$ is an element of Y (Definition 3.2). If $y \in f(S)$,

then we know that $y = f(x)$ for some $x \in S$ by Definition 3.11. Since $S \subseteq X$, we have that $y = f(x)$ for some $x \in X$, which, since $f(x) \in Y$, implies that $y \in Y$). Thus, by Definition 3.12, $f(x) \in f(S)$. By Definition 3.11 and the fact that f is injective, this implies that $x \in S$ (if f were not injective, then it would be possible that $x \notin S$ but $f(x) \in f(S)$, since there could exist some $x' \in S$ such that $f(x') = f(x)$). On the other hand, by Exercise 3.4.2, $S \subseteq f^{-1}(f(S))$. □

7/29: 6. Prove the following lemma. (Hint: start with the set $\{0,1\}^X$ and apply the replacement axiom, replacing each function f with the object $f^{-1}(\{1\})$.) See also Exercise **??**.

Lemma 3.8. *Let X be a set. Then the set*

$$\{Y : Y \text{ is a subset of } X\}$$

is a set.

Proof. We first define a set composed entirely of subsets of X. We then guarantee that this set includes all subsets of X.

Consider the set $\{0,1\}^X$. For any $f \in \{0,1\}^X$, since $\{1\} \subseteq \{0,1\}$ (i.e., is a subset of the range of f), Definition 3.12 allows us to define the (unique) set $f^{-1}(\{1\})$. Now let $P(f, A)$ be the property $A = f^{-1}(\{1\})$. Then for every element $f \in \{0,1\}^X$, there is exactly, i.e., at most one A for which $P(f, A)$ is true. Consequently, by Axiom 3.6, there exists a set $\{A : P(f, A) \text{ is true for some } f \in \{0,1\}^X\}$, or, more briefly, $\{f^{-1}(\{1\}) : f \in \{0,1\}^X\}$. Now by Definition 3.12, each $f^{-1}(\{1\})$ represents the corresponding set $\{x \in X : f(x) \in \{1\}\}$. This implies by Axiom 3.5 that every $f^{-1}(\{1\})$ is a subset of X. Thus, $\{f^{-1}(\{1\}) : f \in \{0,1\}^X\}$ is a set composed entirely of subsets of X.

At this point, all that's left is to guarantee that $\{f^{-1}(\{1\}) : f \in \{0,1\}^X\}$ includes *all* subsets of X, which may be done as follows. Let Z be any subset of X. Then we may define the function $f : X \to \{0,1\}$ by the following.

$$f(x) := \begin{cases} 1 & x \in Z \\ 0 & x \notin Z \end{cases}$$

By Axiom 3.10, $f \in \{0,1\}^X$. Thus, by the first part of this proof, we know that $f^{-1}(\{1\})$ exists and is defined to be equal to $\{x \in X : f(x) \in \{1\}\}$. But Axiom 3.3, $f(x) \in \{1\}$ iff $f(x) = 1$, and by the definition of f, $f(x) = 1$ iff $x \in Z$. Thus, $f^{-1}(\{1\}) = \{x \in X : x \in Z\} = Z$. Since $f^{-1}(\{1\}) \in \{f^{-1}(\{1\}) : f \in \{0,1\}^X\}$, we have $Z \in \{f^{-1}(\{1\}) : f \in \{0,1\}^X\}$. Therefore, there exists a set that contains exactly the subsets of X, which we may denote by $\{Y : Y \text{ is a subset of } X\}$. □

8. Show that Axiom 3.4 can be deduced from Axiom 3.1, Axiom 3.3, and Axiom 3.11.

Proof. To demonstrate Axiom 3.4, we must show that for any two sets A, B, there exists a set $A \cup B$ satisfying $x \in A \cup B$ iff $x \in A$ or $x \in B$.

Let A, B be sets. Then by Axiom 3.1, A and B are objects. Thus, by Axiom 3.3, we can create the pair set $\{A, B\}$. Since $\{A, B\}$ has only sets for elements, Axiom 3.11 applies, and implies the existence of $\bigcup\{A, B\}$, which we may denote by $A \cup B$. Axiom 3.11 also asserts that

$$x \in A \cup B \iff x \in S \text{ for some } S \in \{A, B\}$$

But by Axiom 3.3, $S \in A \cup B$ iff $S = A$ or $S = B$. Therefore, we have

$$x \in A \cup B \iff x \in S,\ S \text{ being a set that satisfies } S = A \text{ or } S = B$$

or

$$x \in A \cup B \iff x \in A \text{ or } x \in B$$

□

9. Show that if β and β' are two elements of a set I, and to each $\alpha \in I$ we assign a set A_α, then

$$\{x \in A_\beta : x \in A_\alpha \text{ for all } \alpha \in I\} = \{x \in A_{\beta'} : x \in A_\alpha \text{ for all } \alpha \in I\}$$

and so the definition of $\bigcap_{\alpha \in I} A_\alpha$ defined in Equation 3.3 does not depend on β. Also explain why Equation 3.4 is true.

Proof. To prove $\{x \in A_\beta : x \in A_\alpha \text{ for all } \alpha \in I\} = \{x \in A_{\beta'} : x \in A_\alpha \text{ for all } \alpha \in I\}$, Definition 3.1 tells us that it will suffice to show that

$$y \in \{x \in A_\beta : x \in A_\alpha \text{ for all } \alpha \in I\} \iff y \in \{x \in A_{\beta'} : x \in A_\alpha \text{ for all } \alpha \in I\}$$

By Equations 3.3 and 3.4, we have

$$y \in \{x \in A_\beta : x \in A_\alpha \text{ for all } \alpha \in I\} \iff y \in A_\alpha \text{ for all } \alpha \in I$$

Similarly, we have

$$y \in \{x \in A_{\beta'} : x \in A_\alpha \text{ for all } \alpha \in I\} \iff y \in A_\alpha \text{ for all } \alpha \in I$$

Therefore, by Exercise A.1.5, we have

$$y \in \{x \in A_\beta : x \in A_\alpha \text{ for all } \alpha \in I\} \iff y \in \{x \in A_{\beta'} : x \in A_\alpha \text{ for all } \alpha \in I\}$$

as desired.

As to the other question, by Definition 3.1, $y \in \bigcap_{\alpha \in I} A_\alpha \implies y \in \{x \in A_\beta : x \in A_\alpha \text{ for all } \alpha \in I\}$. Thus, by Axiom 3.5, $y \in A_\beta$ and $y \in A_\alpha$ for all $\alpha \in I$. But $A_\beta = A_\alpha$ for some $\alpha \in I$, so the condition that $y \in A_\beta$ is a tautology. Thus, $y \in A_\alpha$ for all $\alpha \in I$. A similar argument works in the other direction. □

Appendix A

Appendix: The Basics of Mathematical Logic

6/19:

- **Mathematical logic**: The language one uses to conduct rigorous mathematical proofs.
- "A logical argument may sometimes look unwieldy, excessively complicated, or otherwise appear unconvincing. The big advantage of writing logically, however, is that one can be absolutely sure that your conclusion will be correct" (Tao, 2016, p. 305).
- "Being logical is not the only desirable trait in writing, and in fact sometimes it gets in the way; mathematicians for instance often resort to short informal arguments which are not logically rigorous when they wnat to convince other mathematicians of a statement without going through all of the long details" (Tao, 2016, p. 305).
- "Because logic is innate, the laws of logic that you learn should make sense — if you find yourself having to memorize one of the principles or laws of logic here, without feeling a mental 'click' or comprehending why that law should work, then you will probably not be able to use that law of logic correctly and effectively in practice. So, please don't study this appendix the way you might cram before a final — that is going to be useless. Instead, put away your highlighter pen, and read and understand this appendix rather than merely studying it" (Tao, 2016, p. 306).

A.1 Mathematical Statements

- **Mathematical statement**: A precise statement "concerning various mathematical objects (numbers, vectors, functions, etc.) and relations between them (addition, equality, differentiation, etc.)" (Tao, 2016, p. 306).
  - Mathematical statements are either true or false.
  - For example, $2 + 2 = 4$ is true while $2 + 2 = 5$ is false.
- **Ill-formed** (statement): A statement that is neither true nor false (or perhaps not even a statement). *Also known as* **ill-defined**.
  - For example, $0/0 = 1$.
  - The antonym is **well-formed** or **well-defined** statements.
- A logical argument should include only well-formed statements.
- "It is important, especially when just learning a subject, to take care in keeping statements well-fromed and precise. Once you have more skill and confidence, of course you can afford once again to speak

loosely[1], because you will know what you are doing and won't be in as much danger of veering off into nonsense" (Tao, 2016, p. 307).

- The principle of **proof by contradiction** is "to prove that a statement is true, it suffices to show that it is not false, while to show that a statement is false, it suffices to show that it is not true" (Tao, 2016, p. 307).
- The axioms of logic become more dubious in very non-mathematical situations. One can attempt to apply logic via a mathematical model, but that's science or philosophy. There are other models of logic that attempt to account for these scenarios, but they're beyond the scope of this text.
- Note that statements may be true but not useful ($2 = 2$) or efficient ($4 \leq 4$ — $4 = 4$ would be better), and they may be false but still be useful ($\pi = 22/7$). However, we concern ourselves at present with truth, alone, not usefulness or efficiency as those are to some extent matters of opinion.
- **Expression**: "A sequence of mathematical symbols which produces some mathematical object (a number, matrix, function, set, etc.) as its value" (Tao, 2016, pp. 308–09).
  - For example, $2 + 3 * 5$ is an expression while $2 + 3 * 5 = 17$ is a statement.
  - An expression is neither true nor false, but it can be well- or ill-defined.
- **Relation**: A thing that makes statements out of expressions, such as $=$, $<$, $\geq$, $\in$, $\subset$, etc.
- **Property**: A thing that makes statements out of expressions, such as "is prime," "is continuous," "is invertible," etc[2].
- **Logical connective**: A thing that relates multiple mathematical statements, such as and, or, not, if-then, if-and-only-if, etc.
- **Compound statement**: Two or more mathematical statements joined by logical connectives.
- **Conjunction**: "If X is a statement and Y is a statement, the statement 'X and Y' is true if X and Y are both true, and is false otherwise" (Tao, 2016, p. 309).
  - Logician's notation: "$X \wedge Y$" or "$X\&Y$."
  - Note that "X and Y" can be reworded "X and also Y," or "Both X and Y are true," or even "X, but Y," or a multitude of other ways and convey the same statement logically, if not expressively.
- **Disjunction**: "If X is a statement and Y is a statement, the statement 'X or Y' is true if either X or Y is true, or both" (Tao, 2016, p. 309).
  - This is called **inclusive or**. There is a such thing as **exclusive or**, but it shows up far less regularly.
  - An exclusive or may be, "Either X or Y is true, but not both," or "Exactly one of X or Y is true."
  - To verify a disjunction, it suffices to verify just one case — this comes in handy when it is significantly easier to verify one case over the other.
- **Negation**: "The statement 'X is false'... is true if and only if X is false, and is false if and only if X is true" (Tao, 2016, p. 310).
  - Logician's notation: "$\sim X$," "$!X$," or "$\neg X$."
  - Negations convert "and" into "or": The negation of "Jane has black hair and Jane has blue eyes" is "Jane doesn't have black hair or Jane doesn't have blue eyes." Similarly, the negation of "$x \geq 2$ and $x \leq 6$" is "$x < 2$ or $x > 6$."

[1] This is very similar to the English class/media arts idea of "you have to know the rules to be able to break them."
[2] With the introduction of properties, it is worth officially noting that mathematical statements can contain English words.

  - Negations convert "or" into "and": The negation of "John has black hair or brown hair" is "John does not have black hair and does not have brown hair" or, equivalently, "John has neither black nor brown hair." Similarly, the negation of "$x < -1$ or $x > 1$" is "$x \geq -1$ and $x \leq 1$."
  - Negations can produce clearly false statements: the negation of "x is even or odd" is "x is neither even nor odd."
  - Negations can get unwieldy — be careful!
- Note that "X is true" can be abbreviated to simply "X."
- **If and only if**: "If X is a statement, and Y is a statement, we say that "X is true if and only if Y is true," [if] whenever X is true, Y has to be also, and whenever Y is true, X has to be also (i.e., X and Y are 'equally true')" (Tao, 2016, p. 311). *Also known as* **iff**.
  - Logician's notation: "$X \leftrightarrow Y$."
  - Denotes "logically equivalent statements."
  - For example, "$x = 3$ if and only if $2x = 6$" is true.
  - False statements can be logically equivalent: "$2 + 2 = 5$ if and only if $4 + 4 = 10$."
  - Sometimes, it is of interest to show that more than two statements are logically equivalent (See Exercise A.1.5 and A.1.6).

Exercises

6/22:

1. What is the negation of the statement, "either X is true, or Y is true, but not both?"

   Proof. X and Y or neither X nor Y. □

2. What is the negation of the statement, "X is true if and only if Y is true?" There may be multiple ways to phrase this negation.

   Proof. X is true if and only if Y is false. □

3. Suppose that you have shown that whenever X is true, then Y is true, and whenever X is false, then Y is false. Have you now demonstrated that X and Y are logically equivalent? Explain.

   Proof. Yes — although we have only described the dependence of Y on X, the dependence of X on Y is implied (suppose X is false when Y is true — but since X is false, Y is false, a contradiction; the same holds the other way around). □

4. Suppose that you have shown that whenever X is true, then Y is true, and whenever Y is false, then X is false. Have you now demonstrated that X is true if and only if Y is true? Explain.

   Proof. No — X could be false and Y could still be true without leading to any contradictions. □

5. Suppose you know that X is true if and only if Y is true, and you know that Y is true if and only if Z is true. Is this enough to show that X, Y, Z are all logically equivalent? Explain.

   Proof. Yes — by the supposition, $X \leftrightarrow Y$ and $Y \leftrightarrow Z$. Since X true implies Y true implies Z true, X false implies Y false implies Z false, Z true implies Y true implies X true, and Z false implies Y false implies X false, $X \leftrightarrow Z$. Thus, X, Y, Z are logically equivalent. □

6. Suppose you know that whenever X is true, then Y is true; that whenever Y is true, then Z is true; and whenever Z is true, then X is true. Is this enough to show that X, Y, Z are all logically equivalent? Explain.

   Proof. Yes — clearly X, Y, Z are equally true. If they were not equally false (say if only one or two were false), then we would have a contradiction, as one true implies all the others are true. Thus, they are equally false, too. Therefore, they are logically equivalent. □

A.2 Implication

6/24:

- **Implication**: "If X is a statement, and Y is a statement, then 'if X, then Y' is the implication from X to Y" (Tao, 2016, p. 312).
  - Logician's notation: "$X \implies Y$."
  - Can be reworded "when X is true, Y is true," "X implies Y," "Y is true when X is," or "X is true only if Y is true."
  - "What this statement 'if X, then Y' means depends on whether X is true or false. If X is true, then 'if X, then Y' is true when Y is true, and false when Y is false. If however X is false, then 'if X, then Y' is *always* true, regardless of whether Y is true or false!" (Tao, 2016, p. 312).
  - "When X is true, the statement 'if X, then Y' implies that Y is true. But when X is false, the statement 'if X, then Y' offers no information about whether Y is true or not; the statement is true, but **vacuous**" (Tao, 2016, p. 312).
- **Vacuous** (statement): A statement that "does not convey any new information" (Tao, 2016, p. 312).
  - Vacuous statements can still be helpful in proofs, though.
- The only way to disprove an implication (to show that it's false) is to show that the hypothesis (first statement) is true *and* the conclusion (second statement) is false.
- Whereas "if and only if" asserts that X and Y are equally true, an implication only asserts that Y is at least as true as X.
- Vacuously true implications are often used in a situation where the conclusion and hypothesis are both false, but the implication is true regardless, e.g., "if John had left work at 5pm, then he would be here by now."
- They can also be used to facilitate a proof by contradiction: We know that "if John had left work at 5pm, then he would be here by now" and that "John is not here now." Suppose that John left work at 5pm. Then he would be here by now, a contradiction. Thus, John did not leave work at 5pm.
- Implications can be true even when there is no causal link between the hypothesis and conclusion[3], e.g., if $1 + 1 = 2$, then Washington, D.C. is the capital of the United States is true.
  - Using such acausal implications in a logical argument is typically frowned upon, since it will likely cause unneeded confusion.
- To prove, "if X, then Y," assume X is true and then deduce Y from X and other known hypotheses.
  - It does not matter whether or not X is true; an implication can be proved true irrespective of whether or not X is true.
  - It would be incorrect, though, to assume Y and deduce X. One cannot assume the conclusion and deduce the hypothesis.
- With regard to vacuously true statements, one need not be concerned that some hypotheses in their argument might not be correct, as long as their argument is still structured to give the correct conclusion regardless of whether those hypotheses were true or false (proving $n(n + 1)$ is even when $n \in \mathbb{N}$).
- **Converse** (of "if X, then Y"): The statement "if Y, then X."
  - "X if and only if Y" implies that "if X, then Y" *and* its converse are true.
- **Inverse** (of "if X, then Y"): The statement "if X is false, then Y is false."

[3] This is really weird — I'll need to return to this later.

  - The inverse of a true implication is not necessarily a true implication.
  - If $X \leftrightarrow Y$, then the inverse of "if X, then Y" holds (namely, we know that "if X is false, then Y is false").
- **Contrapositive** (of "if X, then Y"): The statement "if Y is false, then X is false."
  - The contrapositive and the original statement are equally true.
- **Proof by contradiction**: "To show that something must be false, assume first that it is true, and show that this implies something which you know to be false (e.g., that a statement is simultaneoulsy true and not true)" (Tao, 2016, p. 316). *Also known as* **reductio ad absurdum**.
  - Particularly (but not exclusively) useful for proving "negative" statements, e.g., X is false, $a \neq b$, etc.
- On logician's notation: general-purpose mathematicians do not often use them because English words are often more readable and don't take up that much more space. That being said, $\implies$ is a possible exception.

A.3 The Structure of Proofs

6/25:
- **Direct approach**: Assume the hypothesis and work one's way toward a conclusion.
- For example, to prove "A implies B," show A implies C, implies D, implies B. Done.
- We can also work backwards: "To show B, it would suffice to show D. Since C implies D, we just need to show C. But C follows from A" (Tao, 2016, p. 318).
  - Note that this *is* different from starting with the conclusion and proving the hypothesis.
- One can also move forward and backward at will.
  - "Thus there are many ways to write the same proof down; how you do so is up to you, but certain ways of writing proofs are more readable and natural than others, and different arrangements tend to emphasize different parts of the argument" (Tao, 2016, p. 319).
- Multiple hypotheses and conclusions can split the proof into cases, making it more complicated.
- "There are several things to try when attempting a proof. With experience, it will become clearer which approaches are likely to work easily, which ones will probably work but require much effort, and which ones are probably going to fail" (Tao, 2016, p. 319).
- "There may definitely be multiple ways to approach a problem, so if you see more than one way to begin a problem, you can just try whichever one looks the easiest, but be prepared to switch to another approach if it begins to look hopeless" (Tao, 2016, p. 320).

A.4 Variables and Quantifiers

- **Propositional logic**: "Starting with primitive statements (such as '$2 + 2 = 4$' or 'John has black hair'), then forming compound statements using logical connectives, and then using various laws of logic to pass from one's hypotheses to one's conclusions" (Tao, 2016, p. 320). *Also known as* **Boolean logic**.
  - There are a dozen or so laws of propositional logic, but Tao, 2016 excludes them because memorizing such a list is not how one should learn how to do logic. I will, however, include the list (from "The Laws of Propositional Logic", 2018) for reference.

- Let P, Q, and R be statements. Let t be a formula that is a tautology and let f be a formula that is a contradiction. Then:
  1. **Law of Double Negation**: $\neg\neg P \equiv P$.
  2. **Associative Law for Conjunction**: $P \wedge (Q \wedge R) \equiv (P \wedge Q) \wedge R$.
  3. **Associative Law for Disjunction**: $P \vee (Q \vee R) \equiv (P \vee Q) \vee R$.
  4. **Commutative Law for Conjunction**: $P \wedge Q \equiv Q \wedge P$.
  5. **Commutative Law for Disjunction**: $P \vee Q \equiv Q \vee P$.
  6. **First Distributive Law**: $P \vee (Q \wedge R) \equiv (P \vee Q) \wedge (P \vee R)$.
  7. **Second Distributive Law**: $P \wedge (Q \vee R) \equiv (P \wedge Q) \vee (P \wedge R)$.
  8. **Idempotent Law for Conjunction**: $P \wedge P \equiv P$.
  9. **Idempotent Law for Disjunction**: $P \vee P \equiv P$.
  10. **Identity Law for Tautologies**: $P \wedge t \equiv P$.
  11. **Identity Law for Contradictions**: $P \vee f \equiv P$.
  12. **Inverse Law for Tautologies**: $P \vee \neg P \equiv t$.
  13. **Inverse Law for Contradictions**: $P \wedge \neg P \equiv f$.
  14. **Domination Law for Tautologies**: $P \vee t \equiv t$.
  15. **Domination Law for Contradictions**: $P \wedge f \equiv f$.
  16. **De Morgan's Law 1**: $\neg(P \vee Q) \equiv \neg P \wedge \neg Q$.
  17. **De Morgan's Law 2**: $\neg(P \wedge Q) \equiv \neg P \vee \neg Q$.
- Propositional logic is insufficient to do all math, though, because it does not incorporate **variables**.
- **Mathematical logic**: Propositional logic plus **variables**.
- **Variable**: "A symbol, such as n or x, which denotes a certain type of mathematical object — an integer, a vector, a matrix, that kind of thing" (Tao, 2016, p. 320). They denote "quantities which are unknown, or set to some value, or assumed to obey some property" (Tao, 2016, p. 320).
  - In almost all circumstances, the type of object that the variable represents should be declared; variables should have an explicit type.
  - $x = x$ is true regardless of the type, but $x + y = y + x$ is only true if x and y are variables for which addition makes sense and are commutative.
- "The truth of a statement involving a variable may depend on the **context** of the statement" (Tao, 2016, p. 321).
- **Free variable**: A variable with a specified type but an unspecified value.
  - "Statements with free variables might not have a definite truth value" (Tao, 2016, p. 321).
  - $x + 3 = 5$ does not have a definite truth value if x is a free variable.
  - $(x + 1)^2 = x^2 + 2x + 1$ does have a definite truth value even with x as a free variable.
- **Set** (a free variable): Define a value to a free variable via a statement such as "Let $x = 2$" or "Set x equal to 2."
- **Bound variable**: A variable with a specified type that has been set to some value.
  - "Statements involving only bound variables and no free variables do have a definite truth value" (Tao, 2016, p. 321).
- A free variable can be bounded without being set to a *specific* value, per se, with the use of "for all" or "for some."

  - "$(x+1)^2 = x^2 + 2x + 1$" is indefinite; "$(x+1)^2 = x^2 + 2x + 1$ for all $x \in \mathbb{R}$" is definite.
  - "$x + 3 = 5$" is indefinite; "$x + 3 = 5$ for some $x \in \mathbb{R}$" is definite.
- **Universal quantifiers**: A quantifier of the form "for all x of type T."
  - Logician's notation: "$\forall$."
  - "$P(x)$ is true for all x of type T" is equivalent to the implication, "if x is of type T, then $P(x)$ is true."
  - Prove such a statement by letting x be any element of type T and showing that $P(x)$ holds.
  - Disprove it with just a single counterexample (an element $x \in T$ for which $P(x)$ is false).
  - If there are no variables x of type T, then "$P(x)$ is true for all x of type T" is **vacuously true**.
- **Existential quantifiers**: A quantifier of the form "for some x of type T."
  - Logician's notation: "$\exists$."
  - There exists at least one (but possibly numerous) x of type T for which $P(x)$ is true.
    - To indicate uniqueness, use "for exactly one $x\ldots$" instead of "for some $x\ldots$"
  - Prove such a statement by providing a single example of such an x.
  - A vacuously true "for all" statement will be false as a "for some" statement.
    - "$6 < 2x < 4$ for all $3 < x < 2$ is true, but $6 < 2x < 4$ for some $3 < x < 2$ is false" (Tao, 2016, p. 324).

A.5 Nested Quantifiers

6/28:
- Two or more quantifiers can be combined to produce a type of compound statement.
- Negating a universal statement produces an existential statement, and vice versa, as alluded to above.

6/29:
- **Aristotlean logic**: "Deals with objects, their properties, and quantifiers such as 'for all' and 'for some'" (Tao, 2016, p. 326).
  - A subset of mathematical logic since it lacks the concepts of logical connectives and binary relations (e.g., $=$ or $<$).
- **Syllogism**: A typical line of reasoning in Aristotlean logic consisting of a major and minor premise, and a conclusion.
- Swapping the order of quantifiers may or may not make a difference.
  - "For all real numbers a, and for all real numbers b, we have $(a+b)^2 = a^2 + 2ab + b^2$" is logically equivalent to "for all real numbers b, and for all real numbers a, we have $(a+b)^2 = a^2 + 2ab + b^2$."
  - "For every integer n, there exists an integer m which is larger than n" is not logically equivalent to "There exists an integer m which is larger than n for every integer n."
  - "The reason why the order of quantifiers is important is that the inner variables may possibly depend on the outer variables, but not vice versa" (Tao, 2016, p. 327).

Exercises

7/2:
1. What does each of the following statements mean, and which of them are true? Can you find gaming metaphors for each of these statements?
   (a) For every positive number x, and every positive number y, we have $y^2 = x$.

Proof. It means that for all positive numbers x, the statement "$y^2 = x$ for all positive number y" is true. However, this is a false statement ($2^3 \neq 3$, for example). As a gaming metaphor, suppose you play a game where your opponent picks two positive numbers x and y. If $y^2 = x$, then you win. If you can always win, then the statement is true. □

(b) There exists a positive number x such that for every positive number y, we have $y^2 = x$.

Proof. It means that for some positive number x, the statement "$y^2 = x$ for all positive numbers y" is true. However, this is a false statement (no positive number exists such that every positive number squared equals it). As a gaming metaphor, suppose you play a game where you first pick a positive number x, and then your opponent picks a positive number y. If you can pick an x such that any y picked by your opponent satisfies $y^2 = x$, then you win. If you can win once, then the statement is true. □

(c) There exists a positive number x, and there exists a positive number y, such that $y^2 = x$.

Proof. It means that the statement "$y^2 = x$" is true for some pair of positive numbers x, y. This statement is true. As a gaming metaphor, suppose you play a game where you pick two positive numbers x and y. If $y^2 = x$, then you win. If you can win once, then the statement is true. □

(d) For every positive number y, there exists a positive number x such that $y^2 = x$.

Proof. It means that for all positive numbers y, the statement "$y^2 = x$ for some positive number x" is true. This statement is true. As a gaming metaphor, suppose you play a game where your opponent first picks a positive number y, and then you pick a positive number x. If $y^2 = x$, then you win. If you can always win, then the statement is true. □

(e) There exists a positive number y such that for every positive number x, we have $y^2 = x$.

Proof. It means that for some positive number y, the statement "$y^2 = x$ for all positive numbers x" is true. However, this is a false statement (no positive number exists such that the square root of every positive number equals it). As a gaming metaphor, suppose you play a game where you first pick a positive number y, and then your opponent picks a positive number x. If you can pick a y such that any x piked by your opponent satisfies $y^2 = x$, then you win. If you can win once, then the statement is true. □

A.6 Some Examples of Proofs and Quantifiers

- This section proves some simple results involving the "for all" and the "for some" quantifiers to demonstrate how the quantifiers are arranged and how the proofs are structured. Let's begin.
- **Proposition A.1.** *For every $\varepsilon > 0$ there exists a $\delta > 0$ such that $2\delta < \varepsilon$.*

  Proof. Let $\varepsilon > 0$ be arbitrary. We have to show that there exists a $\delta > 0$ such that $2\delta < \varepsilon$. We only need to pick one such δ; choosing $\delta := \frac{\varepsilon}{3}$ will work, since one then has $2\delta = \frac{2}{3}\varepsilon < \varepsilon$. □

  – Since we are proving something for *every* ε, we must make ε arbitrary.
  – On the other hand, δ can be chosen as we wish, because we only need to show that there exists *some* δ which does what we want.
  – The δ can depend on ε because "the δ-quantifier is nested inside the ε-quantifier" (Tao, 2016, p. 328).

- **Proposition A.2.** *There exists an $\varepsilon > 0$ such that $\sin x > \frac{x}{2}$ for all $0 < x < \varepsilon$.*

Proof. We pick $\varepsilon > 0$ to be chosen later, and let $0 < x < \varepsilon$. Since the derivative of $\sin x$ is $\cos x$, we see from the mean-value theorem we have

$$\frac{\sin x}{x} = \frac{\sin x - \sin 0}{x - 0} = \cos y$$

for some $0 \leq y \leq x$. Thus, in order to ensure that $\sin x > \frac{x}{2}$, it would suffice to ensure that $\cos y > \frac{1}{2}$. To do this, it would suffice to ensure that $0 \leq y < \frac{\pi}{3}$ (since the cosine function takes the value of 1 at 0, takes the value of $\frac{1}{2}$ at $\frac{\pi}{3}$, and is decreasing in between). Since $0 \leq y \leq x$ and $0 < x < \varepsilon$, we see that $0 \leq y < \varepsilon$. Thus if we pick $\varepsilon := \frac{\pi}{3}$, then we have $0 \leq y < \frac{\pi}{3}$ as desired, and so we can ensure that $\sin x > \frac{x}{2}$ for all $0 < x < \varepsilon$. □

- Normally, when proving a "there exists (ε) such that X is true" statement, one proceeds by selecting ε carefully, and then showing that X is true for that ε. However, when the selection would require a lot of foresight, we defer it until later in the argument, as above.
- Note that since the value of ε didn't depend on the nested variables x and y, the proof is legitimate.

A.7 Equality

- **Equality**: The most important relation. It links two objects x, y of the same type T. "Given two such objects x and y, the statement $x = y$ may or may not be true; it depends on the value of x and y and also on how equality is defined for the class of objects under consideration" (Tao, 2016, p. 329).
  - For example, $12 \neq 2$ in ordinary arithmetic, but $12 = 2$ in modular arithmetic with modulus 10.
- For the purposes of logic, equality obeys the following four **axioms of equality**.
- **Reflexive axiom**: Given any object x, we have $x = x$.
- **Symmetry axiom**: Given any two objects x and y of the same type, if $x = y$, then $y = x$.
- **Transitive axiom**: Given any three objects x, y, z of the same type, if $x = y$ and $y = z$, then $x = z$.
- **Substitution axiom**: Given any two objects x and y of the same type, if $x = y$, then $f(x) = f(y)$ for all functions or operations f. Similarly, for any property $P(x)$ depending on x, if $x = y$, then $P(x)$ and $P(y)$ are equivalent statements.
  - The first three axioms assert that equality is an **equivalence relation**.
- "Thus, from the point of view of logic, we can define equality on a class of objects however we please, so long as it obeys the reflexive, symmetry, and transitive axioms, and is consistent with all other operations on the class of objects under discussion in the sense that the substitution axiom was true for all of those operations" (Tao, 2016, p. 330).

Exercises

1. Suppose you have four real numbers a, b, c, d and you know that $a = b$ and $c = d$. Use the above four axioms to deduce that $a + d = b + c$.

   Proof. Let $f(x) = x + c$. By the substitution axiom, since $a = b$, $f(a) = f(b) \Rightarrow a + c = b + c$. Let $g(x) = a + x$. By the substitution axiom, since $c = d$, $f(c) = f(d) \Rightarrow a + c = a + d$. By the symmetry axiom, $a + d = a + c$. By the transitive axiom, $a + d = a + c = b + c$. □

A.8 Misc. Notes

- "From a logical point of view, there is no difference between a lemma, proposition, theorem, or corollary — they are all claims waiting to be proved. However, we use these terms to suggest different levels of importance and difficulty. A lemma is an easily proved claim which is helpful for proving other propositions and theorems, but is usually not particularly interesting in its own right. A proposition is a statement which is interesting in its own right, while a theorem is a more important statement than a proposition which says something definitive on the subject, and often takes more effort to prove than a proposition or lemma. A corollary is a quick consequence of a proposition or theorem that was proven recently." (Tao, 2016, p. 25).

References

Jiang, Y. (2020). Solutions to Tao's Analysis I: Chapter 3.1. https://yehongjiang.com/math/solutions-to-taos-analysis-i-chapter-3-1/. Accessed 22 July 2020.

Ojo, S. (2019). Proof of Strong principle of Induction (T. Tao Analysis I). Mathematics Stack Exchange. URL: https://math.stackexchange.com/q/3143086 (version: 2019-08-29).

Tao, T. (2016). *Analysis I* (Third). Texts and Readings in Mathematics. P-19 Green Park Extension, New Delhi 110016, India: Hindustan Book Agency.

Tao, T. (2020). Analysis I errata. https://terrytao.wordpress.com/books/analysis-i/. Accessed 23 July 2020.

The Laws of Propositional Logic. (2018). http://mathonline.wikidot.com/the-laws-of-propositional-logic. Accessed 25 June 2020.