Appendix A

Prerequisites

I'll organize this section better later, but for now, here's a brain dump of some concepts you should be familiar with.

A.1 Basic Set Theory

Definition A.1. A **set** is a collection of objects called **elements**. If A is a set and x is an element of A, we write $x \in A$. Otherwise, we write $x \notin A$.

Definition A.2. The set containing no elements is called the **empty set**, and is denoted by the symbol $\emptyset$.

If we think of a set as a box containing some stuff, then the empty set is a box with nothing in it.

Definition A.3 (Interval Notation)**.** For $a, b \in \mathbb{R}$ with $a < b$, we define the following.

1. $(a, b) = \{x \in \mathbb{R} \mid a < x < b\}$
2. $(a, \infty) = \{x \in \mathbb{R} \mid a < x\}$
3. $(-\infty, b) = \{x \in \mathbb{R} \mid x < b\}$
4. $[a, b] = \{x \in \mathbb{R} \mid a \leq x \leq b\}$

We analogously define $[a, b)$, $(a, b]$, $[a, \infty)$, and $(-\infty, b]$.

Remark A.4. There are a few sets with common names that we should be familiar with.

1. **Natural Numbers:** $\mathbb{N} = \{1, 2, 3, \ldots\}$
2. **Integers:** $\mathbb{Z} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$
3. **Real Numbers:** $\mathbb{R} = (-\infty, \infty)$*

*This is really a cop out. If you look at the definition of the interval $(-\infty, \infty)$, we are being circular.

4. **Complex Numbers:** $\mathbb{C} = \{a + bi \mid a, b \in \mathbb{R}\}$, where $i = \sqrt{-1}$ is the imaginary unit.

Definition A.5. The language associated to sets is specific. We will often define sets using the following notation, called **set builder notation**.

$$S = \{x \in A \mid x \text{ satisfies some condition}\}$$

The first part "$x \in A$" denotes what type of x is being considered. The statements to the right of the colon are the conditions that x must satisfy in order to be members of the set. This notation is read as "The set of all x in A such that x satisfies some condition," where "some condition" is something specific about the restrictions on x relative to A.

Definition A.6. If A and B are sets, then we say that A is a **subset** of B, written $A \subseteq B$, provided that every element of A is also an element of B.

Remark A.7. Observe that $A \subseteq B$ is equivalent to "For all x (in the universe of discourse), if $x \in A$, then $x \in B$." Since we know how to deal with "for all" statements and conditional propositions, we know how to go about proving $A \subseteq B$.

Theorem A.8 (Transitivity of subsets)**.** Suppose that A, B, and C are sets. If $A \subseteq B$ and $B \subseteq C$, then $A \subseteq C$.

Definition A.9. If $A \subseteq B$, then A is called a **proper subset** provided that $A \neq B$. In this case, we may write $A \subset B$ or $A \subsetneq B$.[†]

Definition A.10. Let A and B be sets.

1. The **union** of the sets A and B is $A \cup B = \{x \in U \mid x \in A \text{ or } x \in B\}$.
2. The **intersection** of the sets A and B is $A \cap B = \{x \in U \mid x \in A \text{ and } x \in B\}$.
3. The **set difference** of the sets A and B is $A \setminus B = \{x \in U \mid x \in A \text{ and } x \notin B\}$.
4. The **complement of** A (relative to U) is the set $A^c = U \setminus A = \{x \in U \mid x \notin A\}$.

Definition A.11. If two sets A and B have the property that $A \cap B = \emptyset$, then we say that A and B are **disjoint** sets.

Theorem A.12. Let A and B be sets. If $A \subseteq B$, then $B^c \subseteq A^c$.

Definition A.13. Two sets A and B are **equal** if and only if $A \subseteq B$ and $B \subseteq A$. In this case we write $A = B$.

Remark A.14. Given two sets A and B, if we want to prove $A = B$, then we have to do two separate "mini" proofs: one for $A \subseteq B$ and one for $B \subseteq A$.

Theorem A.15. Let A and B be sets. Then $A \setminus B = A \cap B^c$.

Theorem A.16 (DeMorgan's Law)**.** Let A and B be sets. Then

[†] *Warning:* Some books use $\subset$ to mean $\subseteq$.

1. $(A \cup B)^c = A^c \cap B^c$,
2. $(A \cap B)^c = A^c \cup B^c$.

Theorem A.17 (Distribution of Union and Intersection)**.** Let A, B, and C be sets. Then

1. $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$,
2. $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$.

Definition A.18. Suppose we have a collection $\{A_\alpha\}_{\alpha \in \Delta}$.

1. The **union of the entire collection** is defined via
$$\bigcup_{\alpha \in \Delta} A_\alpha = \{x \mid x \in A_\alpha \text{ for some } \alpha \in \Delta\}.$$
2. The **intersection of the entire collection** is defined via
$$\bigcap_{\alpha \in \Delta} A_\alpha = \{x \mid x \in A_\alpha \text{ for all } \alpha \in \Delta\}.$$

Example A.19. In the special case that $\Delta = \mathbb{N}$, we write

$$\bigcup_{n=1}^{\infty} A_n = \{x \mid x \in A_n \text{ for some } n \in \mathbb{N}\} = A_1 \cup A_2 \cup A_3 \cup \cdots$$

and

$$\bigcap_{n=1}^{\infty} A_n = \{x \mid x \in A_n \text{ for all } n \in \mathbb{N}\} = A_1 \cap A_2 \cap A_3 \cap \cdots$$

Similarly, if $\Delta = \{1, 2, 3, 4\}$, then

$$\bigcup_{n=1}^{4} A_n = A_1 \cup A_2 \cup A_3 \cup A_4$$

and

$$\bigcap_{n=1}^{4} A_n = A_1 \cap A_2 \cap A_3 \cap A_4.$$

Remark A.20. Notice the difference between "$\bigcup$" and "$\cup$" (respectively, "$\bigcap$" and "$\cap$"). The larger versions of the union and intersection symbols very much like the notation that you've likely seen for sums (e.g., $\sum_{i=1}^{\infty} i^2$).

Definition A.21. We say that a collection of sets $\{A_\alpha\}_{\alpha \in \Delta}$ is **pairwise disjoint** if $A_\alpha \cap A_\beta = \emptyset$ whenever $\alpha \neq \beta$.

Exercise A.22. Draw a Venn diagram of a collection of 3 sets that are pairwise disjoint.

Exercise A.23. Provide an example of a collection of three sets, say $\{A_1, A_2, A_3\}$, such that the collection is *not* pairwise disjoint, but

$$\bigcap_{n=1}^{3} A_n = \emptyset.$$

Definition A.24. An **ordered pair** is an object of the form (x, y). Two ordered pairs (x, y) and (a, b) are **equal** if $x = a$ and $y = b$.

Definition A.25. An n-**tuple** is an object of the form $(x_1, x_2, \ldots, x_n)$. Each x_i is referred to as the ith **component**.

Note that an ordered pair is just a 2-tuple.

Definition A.26. If X and Y are sets, the **Cartesian product** of X and Y is defined by

$$X \times Y = \{(x, y) \mid x \in X, y \in Y\}.$$

That is, $X \times Y$ is the set of all ordered pairs where the first element is from X and the second element is from Y. The set $X \times X$ is sometimes denoted by X^2. We similarly define the Cartesian product of n sets, say $X_1, \ldots, X_n$, by

$$\prod_{i=1}^{n} X_i = X_1 \times \cdots \times X_n = \{(x_1, \ldots, x_n) \mid \text{each } x_i \in X_i\}.$$

Exercise A.27. What general conclusion can you make about $X \times Y$ versus $Y \times X$? When will they be equal?

Exercise A.28. If X and Y are both finite sets, then how many elements will $X \times Y$ have? Be as specific as possible.

Exercise A.29. Let $X = [0, 1]$ and let $Y = \{1\}$. Describe geometrically what $X \times Y$, $Y \times X$, $X \times X$, and $Y \times Y$ look like.

A.2 Relations

Definition A.30. Let X and Y be sets. A **relation** from a set X to a set Y is a subset of $X \times Y$. A relation on X is a subset of $X \times X$.

Remark A.31. Different notations for relations are used in different contexts. When talking about relations in the abstract, we indicate that a pair (a, b) is in the relation by some notation like $a \sim b$, which is read "a is related to b."

Remark A.32. We can often represent relations using graphs or digraphs. Given a finite set X and a relation $\sim$ on X, a **digraph** (short for *directed graph*) is a discrete graph having the members of X as vertices and a directed edge from x to y iff $x \sim y$.

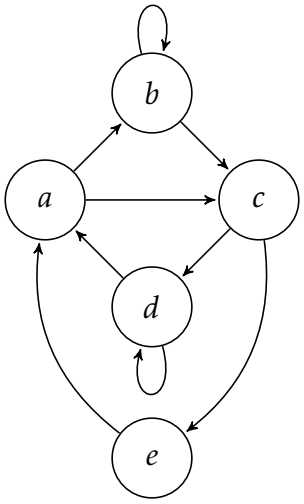

Figure A.1. An example of a digraph for a relation.

Example A.33. Figure A.1 depicts a digraph that represents a relation R given by

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

Exercise A.34. Let $A = \{a,b,c\}$ and define $\sim = \{(a,a),(a,b),(b,c),(c,b),(c,a)\}$. Draw the digraph for $\sim$.

Definition A.35. Let $\sim$ be a relation on a set A.

1. $\sim$ is **reflexive** if for all $x \in A$, $x \sim x$ (every element is related to itself).
2. $\sim$ is **symmetric** if for all $x, y \in A$, if $x \sim y$, then $y \sim x$.
3. $\sim$ is **transitive** if for all $x, y, z \in A$, if $x \sim y$ and $y \sim z$, then $x \sim z$.

Exercise A.36. Given a finite set A and a relation $\sim$, describe what each of reflexive, symmetric, and transitive look like in terms of a digraph.

Exercise A.37. Let P be the set of people at a party and define N via $(x,y) \in N$ iff x knows the name of y. Describe what it would mean for N to be reflexive, symmetric, and transitive.

Definition A.38. Let $\sim$ be a relation on a set A. Then $\sim$ is called an **equivalence relation** if $\sim$ is reflexive, symmetric, and transitive.

Exercise A.39. Determine which of the following are equivalence relations.

1. Let P_f denote the set of all people with accounts on Facebook. Define F via xFy iff x is friends with y.
2. Let P be the set of all people and define H via xHy iff x and y have the same height.

3. Let P be the set of all people and define T via xTy iff x is taller than y.
4. Consider the relation "divides" on $\mathbb{N}$.
5. Let L be the set of lines and define $\|$ via $l_1 \| l_2$ iff l_1 is parallel to l_2.
6. Let $C[0,1]$ be the set of continuous functions on $[0,1]$. Define $f \sim g$ iff
$$\int_0^1 |f(x)|\,dx = \int_0^1 |g(x)|\,dx.$$
7. Define $\sim$ on $\mathbb{N}$ via $n \sim m$ iff $n+m$ is even.
8. Define D on $\mathbb{R}$ via $(x,y) \in D$ iff $x = 2y$.
9. Define $\sim$ on $\mathbb{Z}$ via $a \sim b$ iff $a-b$ is a multiple of 5.
10. Define $\sim$ on $\mathbb{R}^2$ via $(x_1,y_1) \sim (x_2,y_2)$ iff $x_1^2 + y_1^2 = x_2^2 + y_2^2$.
11. Define $\sim$ on $\mathbb{R}$ via $x \sim y$ iff $\lfloor x \rfloor = \lfloor y \rfloor$, where $\lfloor x \rfloor$ is the greatest integer less than or equal to x (e.g., $\lfloor \pi \rfloor = 3$, $\lfloor -1.5 \rfloor = -2$, and $\lfloor 4 \rfloor = 4$).
12. Define $\sim$ on $\mathbb{R}$ via $x \sim y$ iff $|x-y| < 1$.

Definition A.40. Let $\sim$ be a relation on a set A (not necessarily an equivalence relation) and let $x \in A$. Then we define the **set of relatives of** x via
$$[x] = \{y \in A \mid x \sim y\}.$$
Also, define
$$\Omega_\sim = \{[x] \mid x \in A\}.$$
Notice that $\Omega_\sim$ is a set of sets. In particular, an element in $\Omega_\sim$ is a subset of A (equivalently, an element of $\mathcal{P}(A)$). Other common notations for $[x]$ include $\overline{x}$ and R_x.

Exercise A.41. Find $[1]$ and $[2]$ for the relation given in part 9 of Exercise A.39. How many different sets of relatives are there? What are they?

Exercise A.42. If $\sim$ is an equivalence relation on a finite set A, then what is the connection between the equivalence classes and the corresponding digraph?

Theorem A.43. Suppose $\sim$ is an equivalence relation on a set A and let $a, b \in A$. Then $[a] = [b]$ iff $a \sim b$.

Theorem A.44. Suppose $\sim$ is an equivalence relation on a set A. Then

1. $\bigcup_{x \in A} [x] = A$, and
2. for all $x, y \in A$, either $[x] = [y]$ or $[x] \cap [y] = \emptyset$.

Definition A.45. In light of Theorem A.44, if $\sim$ is an equivalence relation on a set A, then we refer to each $[x]$ as the **equivalence class** of x. In this case, $\Omega_\sim$ is the set of equivalence classes determined by $\sim$.

Remark A.46. The upshot of Theorem A.44 is that given an equivalence relation, every element lives in exactly one equivalence class. We'll see in the next section of notes that we can run this in reverse. That is, if we separate out the elements of a set so that every element is an element of exactly one subset (like the bins of my kid's toys), then this determines an equivalence relation. More on this later.

A.3 Partitions

Definition A.47. A collection Ω of nonempty subsets of a set A is said to be a **partition** of A if the elements of Ω satisfy:

1. Given $X, Y \in \Omega$, either $X = Y$ or $X \cap Y = \emptyset$ (We can't have both at the same time. Do you see why?), and
2. $\bigcup_{X \in \Omega} X = A$.

That is, the elements of Ω are pairwise disjoint and their union is all of A.

The next theorem spells out half of the close connection between partitions and equivalence relations. Hopefully you were anticipating this.

Theorem A.48. Let $\sim$ be an equivalence relation on a set A. Then $\Omega_\sim$ forms a partition of A.

Exercise A.49. Consider the equivalence relation

$$\sim = \{(1,1),(1,2),(2,1),(2,2),(3,3),(4,4),(4,5),(5,4),(5,5),(6,6),(5,6),(6,5),(4,6),(6,4)\}$$

on the set $A = \{1,2,3,4,5,6\}$. Find the partition determined by $\Omega_\sim$.

It turns out that we can reverse the situation, as well. That is, given a partition, we can form an equivalence relation. Before proving this, we need a definition.

Definition A.50. Let A be a set and Ω any collection of subsets from $\mathcal{P}(A)$ (not necessarily a partition). If $a, b \in A$, we will define a to be Ω-related to b if there exists an $R \in \Omega$ that contains both a and b. This relation is denoted by $\sim_\Omega$ and is called the **relation on** A **associated to** Ω.

Remark A.51. This definition may look more awkward than the actual underlying concept. The idea is that if two elements are in the same subset, then they are related. For example, when my kids pick up all their toys and put them in the appropriate toy bins, we say that two toys are related if they are in the same bin.

Remark A.52. Notice that we have two notations that looks similar: $\Omega_\sim$ and $\sim_\Omega$.

1. $\Omega_\sim$ is the collection of subsets of A determined by the relation $\sim$.
2. $\sim_\Omega$ is the relation determined by the collection of subsets Ω.

Theorem A.53. Let A be a set and let Ω be a partition of A. Then $\sim_\Omega$ is an equivalence relation.

Remark A.54. The previous theorem says that every partition determines a natural equivalence relation. Namely, two elements are related if they are in the same equivalence class.

A.4 Functions

Definition A.55. Let X and Y be two nonempty sets. A **function** from set X to set Y, denoted $f : X \to Y$, is a relation (i.e., subset of $X \times Y$) such that:

1. For each $x \in X$, there exists $y \in Y$ such that $(x, y) \in f$, and
2. If $(x, y_1), (x, y_2) \in f$, then $y_1 = y_2$.

Note that if $(x, y) \in f$, we usually write $y = f(x)$ and say that "f maps x to y."

Remark A.56. Item 1 of Definition A.55 says that every element of X appears in the first coordinate of an ordered pair in the relation. Item 2 says that each element of X only appears once in the first coordinate of an ordered pair in the relation. It is important to note that there are no restrictions on whether an element of Y ever appears in the second coordinate. Furthermore, if an element of B appears in the second coordinate, it may appear again in a different ordered pair.

Definition A.57. The set X from Definition A.55 is called the **domain** of f and is denoted by $\mathrm{Dom}(f)$. The set Y is called the **codomain** of f and is denoted by $\mathrm{Codom}(f)$. The set

$$\mathrm{Rng}(f) = \{y \in Y \mid \text{there exists } x \text{ such that } y = f(x)\}$$

is called the **range** of f or the **image of** X under f.

Remark A.58. It follows immediately from the definition that $\mathrm{Rng}(f) \subseteq \mathrm{Codom}(f)$. However, it is possible that the range of f is strictly smaller.

Remark A.59. If f is a function and $(x, y) \in f$, then we may refer to x as the **input** of f and y as the **output** of f.

Exercise A.60. Let $X = \{\circ, \square, \triangle, \odot\}$ and $Y = \{a, b, c, d, e\}$. Determine whether each of the following represent functions. Explain. If the relation is a function, determine the domain, codomain, and range.

1. $f : X \to Y$ defined via $f = \{(\circ, a), (\square, b), (\triangle, c), (\odot, d)\}$.
2. $g : X \to Y$ defined via $g = \{(\circ, a), (\square, b), (\triangle, c), (\odot, c)\}$.

3. $h : X \to Y$ defined via $h = \{(\circ, a), (\square, b), (\triangle, c), (\circ, d)\}$.
4. $k : X \to Y$ defined via $k = \{(\circ, a), (\square, b), (\triangle, c), (\odot, d), (\square, e)\}$.
5. $l : X \to Y$ defined via $l = \{(\circ, e), (\square, e), (\triangle, e), (\odot, e)\}$.
6. $m : X \to Y$ defined via $m = \{(\circ, a), (\triangle, b), (\odot, c)\}$.
7. happy : $Y \to X$ defined via happy$(y) = \odot$ for all $y \in Y$.
8. id : $X \to X$ defined via $\mathrm{id}(x) = x$ for all $x \in X$.
9. nugget : $X \to X$ defined via

$$\mathrm{nugget}(x) = \begin{cases} x, & \text{if } x \text{ is a geometric shape,} \\ \square, & \text{otherwise.} \end{cases}$$

Exercise A.61. Let $f : X \to Y$ be a function and suppose that X and Y have n and m elements in them, respectively. Also, suppose that $n < m$. Is it possible for $\mathrm{Rng}(f) = \mathrm{Codom}(f)$? Explain.

Exercise A.62. In high school I am sure that you were told that a graph represents a function if it passes the **vertical line test**. Using our terminology of ordered pairs, explain why this works.

Definition A.63. Two functions are equal if they have the same domain, same codomain, and the same set of ordered pairs in the relation.

Remark A.64. If two functions are defined by the same algebraic formula, but have different domains, then they are *not* equal. For example, the function $f : \mathbb{R} \to \mathbb{R}$ defined via $f(x) = x^2$ is not equal to the function $g : \mathbb{N} \to \mathbb{N}$ defined via $g(x) = x^2$.

Theorem A.65. If $f : X \to Y$ and $g : X \to Y$ are functions, then $f = g$ iff $f(x) = g(x)$ for all $x \in X$.

Definition A.66. Let $f : X \to Y$ be a function.

1. The function f is said to be **one-to-one** (or **injective**) if for all $y \in \mathrm{Rng}(f)$, there is a unique $x \in X$ such that $y = f(x)$.
2. The function f is said to be **onto** (or **surjective**) if for all $y \in Y$, there exists $x \in X$ such that $y = f(x)$.
3. If f is both one-to-one and onto, we say that f is a **one-to-one correspondence** (or a **bijection**).

Exercise A.67. Provide an example of each of the following. You may draw a bubble diagram, write down a list of ordered pairs, or write a formula (as long as the domain and codomain are clear). Assume that X and Y are finite sets.

1. A function $f : X \to Y$ that is one-to-one but not onto.
2. A function $f : X \to Y$ that is onto but not one-to-one.
3. A function $f : X \to Y$ that is both one-to-one and onto.
4. A function $f : X \to Y$ that is neither one-to-one nor onto.

Theorem A.68. Let $f : X \to Y$ be a function. Then f is one-to-one iff for all $x_1, x_2 \in X$, if $f(x_1) = f(x_2)$, then $x_1 = x_2$.

Remark A.69. The previous theorem gives a technique for proving that a given function is one-to-one. Start by assuming that $f(x_1) = f(x_2)$ and then work to show that $x_1 = x_2$.

Remark A.70. To show that a given function is onto, you should start with an arbitrary $y \in \text{Rng}(f)$ and then work to show that there exists $x \in X$ such that $y = f(x)$.

Definition A.71. If $f : X \to Y$ and $g : Y \to Z$ are functions, then a new function $g \circ f : X \to Z$ can be defined by $(g \circ f)(x) = g(f(x))$ for all $x \in \text{Dom}(f)$.

Remark A.72. It is important to notice that the function on the right is the one that "goes first."

Exercise A.73. In each case, give examples of finite sets X, Y, and Z, and functions $f : X \to Y$ and $g : Y \to Z$ that satisfy the given conditions. Drawing bubble diagrams is sufficient.

1. f is onto, but $g \circ f$ is not onto.
2. g is onto, but $g \circ f$ is not onto.
3. f is one-to-one, but $g \circ f$ is not one-to-one.
4. g is one-to-one, but $g \circ f$ is not.

Theorem A.74. If $f : X \to Y$ and $g : Y \to Z$ are both functions that are onto, then $g \circ f$ is also onto.

Theorem A.75. If $f : X \to Y$ and $g : Y \to Z$ are both functions that are one-to-one, then $g \circ f$ is also one-to-one.

Corollary A.76. If $f : X \to Y$ and $g : Y \to Z$ are both one-to-one correspondences, then $g \circ f$ is also a one-to-one correspondence.

Definition A.77. Let $f : X \to Y$ be a function. The relation f^{-1}, called f **inverse**, is defined via

$$f^{-1} = \{(f(x), x) \mid x \in X\}.$$

Remark A.78. Notice that we called f^{-1} a relation and not a function. In some circumstances f^{-1} will be a function and sometimes it won't be.

Exercise A.79. Provide an example of a function $f : X \to Y$ such that f^{-1} is *not* a function. A bubble diagram is sufficient.

Theorem A.80. Let $f : X \to Y$ be a function. Then f^{-1} is a function iff f is 1-1.

Theorem A.81. Let $f : X \to Y$ be a function and suppose that f^{-1} is a function. Then

1. $(f \circ f^{-1})(x) = x$ for all $x \in Y$, and
2. $(f^{-1} \circ f)(x) = x$ for all $x \in X$.

(You only need to prove one of these statements; the other is similar.)

Theorem A.82. Let $f : X \to Y$ and $g : Y \to X$ be functions such that f is a one-to-one correspondence. If $(f \circ g)(x) = x$ for all $x \in Y$ and $(g \circ f)(x) = x$ for all $x \in X$, then $g = f^{-1}$.

Remark A.83. The upshot of the previous two theorems is that if f^{-1} is a function, then it is the only one satisfying the two-sided "undoing" property exhibited in Theorem A.81.

The next theorem can be considered to be a converse of Theorem A.82.

Theorem A.84. Let $f : X \to Y$ and $g : Y \to X$ be functions satisfying $(f \circ g)(x) = x$ for all $x \in Y$ and $(g \circ f)(x) = x$ for all $x \in X$. Then f is a one-to-one correspondence.

Theorem A.85. Let $f : X \to Y$ and $g : Y \to Z$ be functions. If f and g are both one-to-one correspondences, then $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.

A.5 Induction

Induction is a technique for proving statements of the form "For all $n \in \mathbb{N}$, $P(n)$," where $P(n)$ is some predicate involving n. Notice that this is a statement about natural numbers and not some other set.

Axiom A.86 (Axiom of Induction)**.** Let $S \subseteq \mathbb{N}$ such that both

1. $1 \in S$, and
2. if $k \in S$, then $k + 1 \in S$.

Then $S = \mathbb{N}$.

Remark A.87. Recall that an axiom is a basic mathematical assumption. That is, we are assuming that the Axiom of Induction is true, which I'm hoping that you can agree is a pretty reasonable assumption. I like to think of the first hypothesis of the Axiom of Induction as saying that we have a first rung of a ladder. The second hypothesis says that if we have some random rung, we can always get to the next rung. Taken together, this says that we can get from the first rung to the second, from the second to the third, and so on. Again, we are assuming that the "and so on" works as expected here.

Theorem A.88 (Principle of Mathematical Induction)**.** Let $P(1), P(2), P(3), \ldots$ be a sequence of statements, one for each natural number.[‡] Assume

1. $P(1)$ is true, and
2. If $P(k)$ is true, then $P(k+1)$ is true.

Then $P(n)$ is true for all $n \in \mathbb{N}$.[§]

Remark A.89. The Principal of Mathematical Induction (PMI) provides us with a process for proving statements of the form: "For all $n \in \mathbb{N}$, $P(n)$," where $P(n)$ is some predicate involving n. Hypothesis (1) above is called the **base step** while (2) is called the **inductive step**.

Skeleton Proof A.90 (Proof by induction for $(\forall n \in \mathbb{N})P(n)$)**.** Here is what the general structure for a proof by induction looks like. Remarks are in parentheses.

Proof. We proceed by induction.

(i) Base step: (Verify that $P(1)$ is true. This often, but not always, amounts to plugging $n = 1$ into two sides of some claimed equation and verifying that both sides are actually equal. Don't assume that they are equal!)

(ii) Inductive step: (Your goal is to prove that "For all $k \in \mathbb{N}$, if $P(k)$ is true, then $P(k+1)$ is true.") Let $k \in \mathbb{N}$ and assume that $P(k)$ is true. (Now, do some stuff to show that $P(k+1)$ is true.) Therefore, $P(k+1)$ is true.

Thus, by the PMI, $P(n)$ is true for all $n \in \mathbb{N}$. □

[‡]In this case, you should think of $P(n)$ as a predicate, where $P(1)$ is the statement that corresponds to substituting in the value 1 for n.

[§]*Hint:* Let $S = \{k \in \mathbb{N} \mid P_k \text{ is true}\}$ and use the Axiom of Induction. The set S is sometimes called the *truth set*. Your job is to show that the truth set is all of $\mathbb{N}$.