

CHAPTER 1: THE PEANO AXIOMS

WAYNE AITKEN AND LINDA HOLT

Summer 2019 Edition

1. INTRODUCTION

We begin our exploration of number systems with the most basic number system: the natural numbers \mathbb{N} . Informally, natural numbers are just the ordinary whole numbers $0, 1, 2, \dots$ starting with 0 and continuing indefinitely.¹ For a formal description, see the axiom system presented in the next section.

Throughout your life you have acquired a substantial amount of knowledge about these numbers, but do you know the reasons behind your knowledge? Why is addition commutative? Why is multiplication associative? Why does the distributive law hold? Why is it that when you count a finite set you get the same answer regardless of the order in which you count the elements? In this and following chapters we will systematically prove basic facts about the natural numbers in an effort to answer these kinds of questions. Sometimes we will encounter more than one answer, each yielding its own insights. You might see an informal explanation and then a formal explanation, or perhaps you will see more than one formal explanation. For instance, there will be a proof for the commutative law of addition in Chapter 1 using induction, and then a more insightful proof in Chapter 3 involving the counting of finite sets.

We will use the axiomatic method where we start with a few axioms and build up the theory of the number systems by proving that each new result follows from earlier results. In the first few chapters of these notes there will be a strong temptation to use unproved facts about arithmetic and numbers that are so familiar to us that they are practically part of our mental DNA. Resist this temptation! In the context of a formal proof, take the attitude that such familiar facts are not certain until they are proved. So they cannot be used in a formal proof until after they have been proved. A similar thing can be said of definitions: pretend that your intuitive ideas of even basic things such as $+$ and $<$ are inaccessible until you can have a formal definition. In the beginning, the only terms that can be used are

Copyright ©2007–2019 by Wayne Aitken and Linda Holt. The copyright holders authorize individuals to make a single paper copy of this edition for personal, noncommercial use.

¹Warning: some authors do not include 0 in the set of natural numbers. This will be discussed in the next section.

terms from logic and set theory, explained in Chapter 0, and the primitive terms. The only facts that can be used are the axioms together with facts from logic and set theory as summarized in Chapter 0, including general facts about equality, functions, and relations.²

The system of axioms we use here is a famous system called the *Dedekind-Peano axioms* (Section 2), or the *Peano axioms* for short. We will add to this an axiom about iterating functions (Section 3), but in an optional section (Section 9) to this chapter, we will see that this iteration axiom is not necessary since it can actually be proved from Peano's axioms. Thus it is strictly speaking a convenient "temporary" axiom: one could replace the iteration axiom by a theorem that says the same thing. We take it as a temporary axiom in these notes since the proof of the iteration axiom is a bit subtle, and is at a higher level than most of the other theorems of the chapter. I do not want to start off the chapter by scaring away readers.

Remark 1. Although we will be strict about not using unproved assertions in the formal development, you do not need to be so shy about using your prior knowledge in the *informal* exercises. Such prior knowledge is also useful for temporarily guiding your thinking until a firmer foundation is laid down in the formal development.

This distinction between formal and informal is especially important in the many exercises that will arise in these notes. The informal exercises will be labeled as such. The rest are considered to be formal exercises.

The formal exercises may require you to fill in details of sketchy proofs or even to write complete proofs for theorems whose proofs are not too hard or are similar to earlier proofs. These constitute part of the official development of the number systems, and the facts established in them can be used in future proofs. On the other hand, the informal exercises are designed to help familiarize you with facts or definitions, or to lead you in interesting but tangential directions. These do not have to be solved with a formal proof, and can appeal to prior knowledge. They are considered to be outside the logical development of the number systems, and so cannot be cited in a later formal proof.

For example, suppose an informal exercise asks for an example of an associative binary operation that is not commutative. Suppose you know about matrix multiplication from a linear algebra course. Then you can use your knowledge of linear algebra to help solve the problem. On the other hand, you cannot use matrix multiplication in a formal exercise since matrices are not developed in this course.

²In these notes, we start almost at the very beginning of mathematics, but you should be aware that there are other approaches that start with less and begin by proving theorems about set theory first before developing the number systems. For example, set theorists typically start with the Zermelo-Fraenkel axioms for set theory, and from there develop set theory, the number systems, and (most of) the rest of mathematics.

Remark 2. In the above discussion, the term *theorem* refers to any result that has a proof. Keep in mind that other terms for theorems are commonly used including *proposition*, *lemma*, and *corollary*. The term *lemma* is used for a theorem that is only important as a stepping stone in proving other theorems, and a *corollary* is a theorem that follows fairly easily, for example as an interesting special case, from a previous theorem. Some authors also make a distinction between the terms *theorem* and *proposition*, using the label *proposition* for more ordinary theorems and using *theorem* only for the more important theorems. These are informal guidelines: one can find exceptions.

Remark 3. As mentioned above, in the formal development of the natural numbers we begin by assuming that everything about the natural numbers is as yet unknown territory. On the other hand, we do allow logic as expressed in everyday, but careful, language. This leads to a point that needs to be clarified: even though we are developing the natural numbers from scratch, we will allow ourselves to use a few number-related terms such as “pair”, “unique”, “first”, “second”, and so on. We do so because we can safely treat such basic terms as forming part of our *logical* vocabulary.³ We will also use numerals for the labeling of sections, theorems, exercises, and such. These labels have no arithmetic content, and could have just as easily been any string of symbols. They are being used informally to help keep the chapter organized. On the other hand, we will not take any truly mathematical or arithmetic fact for granted, for example facts about addition and multiplication. These all must be proved.

2. THE AXIOMS

Forget everything you think you know about the natural numbers, even something as basic as $1 + 1 = 2$. Pretend you don't even know the definition of addition. In what follows, we will recreate all this knowledge on a solid logical foundation by *proving* all the elementary theorems and *defining* all the basic ideas. (Of course this self-imposed forgetting should be confined to the official formal development of the natural numbers, and the formal proofs. Your past knowledge will come in handy for thinking up strategies for proofs, for helping you mentally digest definitions, and for warning you when you are about to make an error.)

At this point, the only thing that you are officially allowed to know concerning the natural numbers is what is expressed in the following axioms. They function partially as descriptions of the primitive terms, and partially as a list of facts that we can use in later proofs. These axioms are called

³For example, the statement “the set S has at least two elements” does not really require the number 2. It can be translated easily into basic logic as follows:

$$\exists x \exists y \left((x \in S) \wedge (y \in S) \wedge (x \neq y) \right).$$

the *Dedekind-Peano axioms* since they are based on the axioms of the German mathematician Richard Dedekind (1831 – 1916) and the the Italian mathematician Giuseppe Peano (1858 – 1932).⁴

We begin with the primitive terms described in the axioms. They are called *primitive* because they do not have to be formally defined, but instead are described in the axioms. All other terms, such as $+$ or $<$ must be defined. Such definitions can build on primitive terms, notions from Chapter 0, or any previously defined term.

Primitive Terms. *The three primitive terms are \mathbb{N} , 0 , and σ .*

The axioms then tell you everything you are allowed to assume about the meaning of these terms. For example, the first axiom tells you broadly what type of object these terms denote.

Axiom 1. *(i) \mathbb{N} is a set, (ii) 0 is an element of \mathbb{N} , and (iii) σ is a function*

$$\sigma : \mathbb{N} \rightarrow \mathbb{N}$$

with domain and codomain equal to \mathbb{N} .

We call \mathbb{N} the “set of natural numbers”, and we call its elements “natural numbers”. We call 0 the “zero element”, or just “zero”. We call σ the “successor function”. If $n \in \mathbb{N}$ we call σn the “successor of n .” Informally, the successor of n is the next number following n . This is informal since we have not yet defined an order $<$ on \mathbb{N} .

Axiom 2. *The image of $\sigma : \mathbb{N} \rightarrow \mathbb{N}$ does not contain 0 :*

$$\neg(\exists n \in \mathbb{N}, \sigma n = 0).$$

In other words, 0 is not the successor of a natural number.

Axiom 3. *The function $\sigma : \mathbb{N} \rightarrow \mathbb{N}$ is injective.⁵ In other words, distinct natural numbers have distinct successors.*

$$\forall x, y \in \mathbb{N}, x \neq y \implies \sigma x \neq \sigma y$$

or equivalently

$$\forall x, y \in \mathbb{N}, \sigma x = \sigma y \implies x = y.$$

⁴There are several variations of these axioms. We use a version of what is sometimes called the *second-order Peano axioms* which allows the notion of subsets of \mathbb{N} . There is another, more elementary system called the *first-order Peano axioms* which does not quantify over sets of natural numbers. If you encounter the Peano axioms outside these notes, you might see the first order version with axioms that refer directly to addition and multiplication. In our second-order version the operations of addition and multiplication are not mentioned in the axioms, but must be defined in terms of the successor function.

⁵The reader is expected to be familiar with the term *injective*, or the equivalent term *one-to-one*. These terms describe functions f that map distinct elements to distinct images.

Axiom 4 (Induction). *Suppose S is a subset of \mathbb{N} such that (i) $0 \in S$, and (ii) $n \in S$ implies $\sigma n \in S$ for arbitrary $n \in \mathbb{N}$. Then $S = \mathbb{N}$.*

$$S \subseteq \mathbb{N} \wedge 0 \in S \wedge \left(\forall n (n \in S \Rightarrow \sigma n \in S) \right) \implies S = \mathbb{N}$$

Informal Exercise 1. Go through the axioms one by one, and convince yourself that they do indeed hold for your conception of the natural numbers

$$0, 1, 2, 3, \dots$$

Informally think of σn as the next number after n , or as $n + 1$. Since the exercise is informal, you may appeal to your earlier knowledge of arithmetic, knowledge of which will be formally proved later in the course. If it helps to justify the Induction Axiom (Axiom 4), think about why there could not be a smallest natural number $m \notin S$ given (i) and (ii) are known for $S \subseteq \mathbb{N}$.

Remark 4. As discussed in the introduction, basic concepts related to logic, sets, functions, and equality are all taken as given. They constitute the *logical background* to the development while the Dedekind-Peano axioms above are what we take to be the first real *mathematical* assumptions.⁶

Remark 5. Since \mathbb{N} , 0 , and σ are primitive, they do not have to be defined. We start with some undefined terms to avoid circular definitions. All that we know about these terms at the moment is what is set forth in the axioms. By taking \mathbb{N} to be primitive, we are avoiding the question “what are the natural numbers really”. Our answer is just that they are elements of \mathbb{N} where \mathbb{N} is some set satisfying the axioms. Mathematicians regard the question “what are the natural numbers really” as not a mathematical question but as a philosophical question. Such questions have actually played an important role in the history of philosophy for thousands of years, and continue to be discussed in contemporary philosophy.

Remark 6. Some authors, especially of older texts, view the natural numbers as starting with 1. The axioms are then written in terms of 1 instead of 0. It makes sense to begin with 1 from a historical point of view since it took many years for mathematicians to get comfortable with the number 0. So in some sense 0 is not as natural as the positive integers. On the other hand, one of the main reasons for developing the natural numbers is for counting the size, or cardinality, of finite sets. Today the empty set \emptyset is in common use, and we need 0 to describe its cardinality.

Remark 7. You might have seen induction presented in a slightly different style than that given above. Our axiom is in terms of *sets*, but you may have seen induction described in terms of *properties* instead. Perhaps it was stated in terms of identifying a certain property or statement that you want to prove for all of \mathbb{N} by (i) proving it for 0 (the base case) and (ii) assuming

⁶The real location of the line between logic and mathematics is an interesting philosophical issue with no one predominate answer. The line drawn here is convenient for our purposes.

it for n (the inductive hypothesis) and then proving it for the next number after n (which we call σn , but it is commonly called $n + 1$).

The two versions of induction, however, are really the same. To see why, we need to think about the connection between properties and sets. A basic fact of set theory is that every property of natural numbers defines a subset of \mathbb{N} . Showing that a property holds for all natural numbers is the same as showing that the corresponding subset S is all of \mathbb{N} . To see an example, consider the following problem from number theory (using ideas we haven't defined formally yet). Suppose you want to prove that every natural number is the sum of four squares. Then instead of using the property “ n is the sum of four squares” for an inductive hypothesis, our version of induction (Axiom 4) would use the set

$$S = \{n \in \mathbb{N} \mid n \text{ is the sum of four squares}\}.$$

The base step is to show $0 \in S$. By the definition of S this actually amounts to showing that 0 is the sum of four squares ($0 = 0^2 + 0^2 + 0^2 + 0^2$), so it amounts to the same thing as the base case of the other form of induction. Next you need to establish (ii) by assuming $n \in S$ and showing $\sigma n \in S$. By definition of S this means that you assume that n is the sum of four squares (the inductive hypothesis), and somehow try to show that the successor σn is also the sum of four squares (this is the hard part of the proof). Once the base step (i) and the inductive step (ii) have been established, the induction axiom shows that $S = \mathbb{N}$. In other words, all natural numbers are the sum of four squares. As this illustrates, using a set instead of a property is just a very minor change of outlook, the actual work is the same.

In a later chapter we will discuss another type of induction, strong induction, which is truly different from that described above. We will also discuss versions where 0 is replaced by other “base cases”. Our first version of induction is an axiom (Axiom 4), but the later versions will be theorems. The later versions cannot be used until they are proved.

Remark 8. The induction axiom is more complicated than the others. There is a cleaner way of stating this axiom using the notion of “closed” which we now explain. If A is a subset of \mathbb{N} then the image set $\sigma[A]$ is necessarily also a subset of \mathbb{N} since σ is a function $\mathbb{N} \rightarrow \mathbb{N}$. The subset $A \subseteq \mathbb{N}$ is said to be *closed under successor* if $\sigma[A] \subseteq A$. In other words, σ cannot move you out of A : for all $n \in A$, we have $\sigma n \in A$.

Using this concept, we can express the axiom as follows:

If $A \subseteq \mathbb{N}$ contains 0 and is closed under successor then $A = \mathbb{N}$.

Informal Exercise 2. Describe three distinct subsets of \mathbb{N} that are closed under σ but that are not all of \mathbb{N} . By the above remark, none of your examples can contain 0. This shows the importance of checking the “base case” since all of these satisfy (ii) but not (i) of Axiom 4. Hint: since this is informal you have available the formula $\sigma n = n + 1$ even though it has not been proved yet. Also, one of your examples can be the empty set.

We are ready for our first formal definition.

Definition 1. Define 1 as $\sigma 0$. Define 2 as $\sigma 1$.

Exercise 3. Give formal definitions of 3, 4, 5, 6, 7, 8, 9. Now we have names for at least a few numbers. We will wait until Chapter 5 before we develop the familiar base ten notation for naming the rest of the natural numbers.

Remark 9. Symbolic names for numbers are called *numerals*. There is a difference between numbers and numerals since several names can refer to the same number. Suitably defined, IV and 4 refer to the same number: $\sigma(\sigma(\sigma(\sigma 0)))$. So ‘IV’ and ‘4’ are two different names, or numerals, for the same number.⁷

Informal Definition 2 (Stroke numerals). One can regard any string of strokes | as a numeral via the following rule: replace each stroke with σ and then apply these to 0. For example, ||||| is just $\sigma(\sigma(\sigma(\sigma(\sigma 0))))$. Thus ‘|||||’ and ‘6’ are two numerals, from two different systems, for the same number. The Babylonians and the Egyptians used stroke numerals for numbers up to nine, so would have used something like ‘|||||’ for six (but on two rows). The Romans would have used ‘VI’ for the same number. The stroke numerals gives a way of naming all natural numbers. However, this convention is not very efficient or practical! Fortunately we will later develop the more efficient base ten positional notation.

We now end this section by using the axioms to study the concept of *predecessor*. While the successor was primitive and did not have to be defined, predecessor needs to be defined. It is defined in terms of σ :

Definition 3 (Predecessor). Suppose $a, b \in \mathbb{N}$. We say that “ a is a predecessor of b ” if $\sigma a = b$. We say that “ b has a predecessor in \mathbb{N} ” if⁸ there exists an $x \in \mathbb{N}$ such that x is a predecessor of b .

Next we see the first official theorem of the course. It is a simple proof by contradiction.

Theorem 1. *The natural number 0 does not have a predecessor in \mathbb{N} .*

Proof. Suppose otherwise that 0 has predecessor $x \in \mathbb{N}$. By Definition 3 we have $\sigma x = 0$. This contradicts Axiom 2. Thus 0 has no predecessor in \mathbb{N} . \square

⁷A random person on the street might think of numerals and numbers as the same thing. But numerals are symbols. If numbers are not symbols, what are they? This comes back to the philosophical question: *what are numbers really?* As mentioned above we sidestep this as follows: numbers are what the axioms postulate to exist. The axioms do not specify what they really are, they just specify some of their properties. Numerals, on the other hand, are names we give to the objects described by the axioms. In summary, the axioms supply the numbers, but we supply the numerals to refer to these numbers, and can do so any way we choose

⁸In *definitions* “if” really means “if and only if” in common mathematical writing.

Now we see the first proof by induction. It is subtle in one respect. One might want S to be the set $\{x \in \mathbb{N} \mid x \text{ has a predecessor in } \mathbb{N}\}$ for the induction, but this definition of S does not contain 0. So Axiom 4 cannot be used! We do not yet have a form of induction that starts at 1 (we will establish such an induction later). So instead we just artificially put 0 in S by using $\{x \in \mathbb{N} \mid (x = 0) \vee (x \text{ has a predecessor in } \mathbb{N})\}$. We only use this trick when we want to prove something about everything but 0.

Theorem 2. *Every nonzero element of \mathbb{N} has a predecessor in \mathbb{N} .*

Proof. Our goal is to use the induction axiom (Axiom 4). To do so we need to define a set:

$$S \stackrel{\text{def}}{=} \{x \in \mathbb{N} \mid (x = 0) \vee (x \text{ has a predecessor in } \mathbb{N})\}.$$

Observe (i) $0 \in S$ by definition of S .

Next we will establish that (ii) $n \in S \implies \sigma n \in S$ for all $n \in \mathbb{N}$. So assume $n \in S$. Since \mathbb{N} is the codomain of σ we have $\sigma n \in \mathbb{N}$. Observe that n is a predecessor of σn by Definition 3, so σn has a predecessor. Thus $\sigma n \in S$ by definition of S .

Now that we have established (i) and (ii) above, we can use Axiom 4 to conclude that $S = \mathbb{N}$. Since $S = \mathbb{N}$, every element of \mathbb{N} is either 0 or has a predecessor in \mathbb{N} . So if $n \in \mathbb{N}$ and $n \neq 0$ we have that n has a predecessor in \mathbb{N} . \square

We now consider the question of uniqueness. Successors are unique simply because they are values of a function. On the other hand we did not define predecessors as values of a function. Note $\exists!$ denotes “there exists a unique.”⁹

Exercise 4. Prove that if $n \in \mathbb{N}$ has a predecessor in \mathbb{N} then the predecessor is unique. In other words, show the following for all $b \neq 0$ in \mathbb{N} :

$$\exists! a \in \mathbb{N}, a \text{ is a predecessor of } b.$$

Hint: use Axiom 3.

Definition 4 (Positive). A *positive natural number* is a nonzero element of \mathbb{N} . Let \mathbb{N}^+ be the set of positive natural numbers.

The following is an immediate consequence of the above theorem, exercise, and definition.

Corollary 3. *If $n \in \mathbb{N}^+$ then n has a unique predecessor in \mathbb{N} .*

Definition 5 (Predecessor function). We define the *predecessor function* $\pi : \mathbb{N}^+ \rightarrow \mathbb{N}$ as follows: given $n \in \mathbb{N}^+$ we define πn to be the unique predecessor of n .

⁹We do not use ! by itself to mean “unique”. The use of the exclamation mark to mean “unique” is only used after \exists .

In one sense the predecessor function and the successor function are inverses since one undoes the effect of the other. However this cannot be literally true. Since π is a function $\mathbb{N}^+ \rightarrow \mathbb{N}$, its inverse (if it exists) must be a function $\mathbb{N} \rightarrow \mathbb{N}^+$. To deal with this technicality we define a *modified successor function*.

Definition 6 (Modified successor). We define the modified *successor function* $\sigma' : \mathbb{N} \rightarrow \mathbb{N}^+$ as follows: Given $n \in \mathbb{N}$ we define $\sigma'n$ to be σn . Since σn is not 0 (Axiom 2), and since $\sigma'n$ is just σn , we know that $\sigma'n$ is in the set \mathbb{N}^+ . So this definition yields a function with codomain \mathbb{N}^+ as desired.

Observe that $\sigma'n = \sigma n$ for all $n \in \mathbb{N}$. The only difference between the functions is the codomain.

Exercise 5. Let $a \in \mathbb{N}$ and $b \in \mathbb{N}^+$. Show that $\pi b = a$ if and only if $\sigma'a = b$.

Exercise 6. Let $a \in \mathbb{N}$. Show that $\pi(\sigma'a) = a$. Hint: let $b = \sigma'a$ and substitute for b in the above exercise.

Exercise 7. Let $b \in \mathbb{N}^+$. Show that $\sigma'(\pi b) = b$.

Exercise 8. Show that π and σ' are inverse functions. Conclude that they are both bijections.

Hint: recall that $f : A \rightarrow B$ and $g : B \rightarrow A$ are called *inverse functions* if (i) $g(f(x)) = x$ for all $x \in A$, and (ii) $f(g(y)) = y$ for all $y \in B$. Recall also that a function is a bijection if it is both injective and surjective. Finally, recall that a function $f : A \rightarrow B$ is bijective if and only if it has an inverse function (from B to A).

Exercise 9. We know that σ' is bijective. Show that σ is not a bijection.

Exercise 10. The mathematician Dedekind defined a set S to be *infinite* if there is a bijection $S \rightarrow T$ where T is a proper subset of S . Explain why \mathbb{N} is infinite according to Dedekind's definition.¹⁰ (We will give another definition of infinite in Chapter 3).

Exercise 11. Show that if $n \in \mathbb{N}$ then $n \neq \sigma n$. Do so by defining a certain set $S \subseteq \mathbb{N}$ and using the induction axiom to show $S = \mathbb{N}$.

In particular this shows that $0 \neq 1$, and $1 \neq 2$, and so on. It does not mean $0 \neq 2$ though, this has to be proved separately!

3. ITERATION

At this point the only operations we have are successor and predecessor. But any self-respecting theory of arithmetic also needs addition and multiplication. Our strategy for developing these operations is simple: we define addition in terms of iterated successor, and multiplication in terms of iterated addition. Continuing on, we will define exponentiation in terms of iterated multiplication. These definitions all rely on the general concept of *iteration*, so in order to reach our goal of basic arithmetic, we need to take a side trip through iteration.

¹⁰A *proper* subset of S is a subset that is not equal to S .

Informally, we can think of iteration in terms of repeating an action or processes.¹¹ In these notes we think of operations, actions, processes, and such in terms of functions. So iteration will mean repeatedly applying a function.

For example, applying the function $f : S \rightarrow S$ twice to an element $x \in S$ yields $f(f(x))$, which is the same as applying the composition $f \circ f$ once to x . Likewise, applying $f \circ f \circ f$ to x gives the third iteration, and so on. We see that there is a close relationship between repeated composition and iteration. Note that in order to be able to compose a function with itself, it must have a codomain that matches its domain. So we want f to be a function $S \rightarrow S$ for some set S . In summary:

Informal Definition 7. Let $f : S \rightarrow S$ be a function. Observe that we are restricting ourselves to a function whose domain and codomain agree. The *second iteration* f^2 is $f \circ f$, the *third iteration* f^3 is $f \circ f \circ f$, the *fourth iteration* f^4 is $f \circ f \circ f \circ f$, and so on. In general, if $n \geq 2$ is a natural number, the *n*th iteration f^n is obtained by composing f with itself n times.

Remark 10. This is just an informal definition because some ideas in it, such as “composing f with itself n times”, have not been formally defined.

Remark 11. We assume that the reader is already knowledgeable about composition of functions (Chapter 0). Recall that $f \circ g$ is only defined if the codomain of g is equal to the domain of f . Another important fact: composition is associative (when it is defined):

$$f \circ (g \circ h) = (f \circ g) \circ h.$$

This fact allows us to drop parenthesis without introducing ambiguity. So $f \circ g \circ h$ can refer to either $f \circ (g \circ h)$ or $(f \circ g) \circ h$, but both possibilities are equal by the associative law for composition.

What is f^n if n is 0 or 1? Informally, it makes sense to define f^1 as f itself since if you apply this function 1 time to an x in the domain, you get $f(x)$. What if you apply f to x zero times? You will just have x . So it makes sense, informally speaking, to define f^0 as the identity function.

Here is the formal axiom:

Axiom 5 (Iteration). *Let $f : S \rightarrow S$ be a function from a set S to itself, and $n \in \mathbb{N}$. Then the n th iteration of f is a function from S to itself. We write the n th iteration of f as $f^n : S \rightarrow S$. Such functions satisfy the following: (i) f^0 is the identity function on S , and (ii) $f^{\sigma n} = f \circ f^n$.*

Remark 12. Here iteration is regarded as a primitive notion. In a later section (Section 9), however, we will see that there is a way in which the n th iteration can be *defined* and the the properties (i) and (ii) *proved*. Thus, for those willing to do some extra work, the above can be converted from an axiom to a theorem.

¹¹The verb *iterate* comes from the Latin verb *itero* meaning ‘repeat’.

I have decided to move the proof to Section 9 because it is fairly long and a bit tricky, and because I want to get to basic arithmetic as soon as possible. Ultimately, however, it is an “eliminatable axiom”.

Exercise 12. Use this axiom to prove that

$$f^1 = f, \quad f^2 = f \circ f, \quad f^3 = f \circ (f \circ f).$$

Informal Exercise 13. Consider the function $f : \mathbb{R} \rightarrow \mathbb{R}$ defined by the rule $x \mapsto 3x$. Give a formula for the fifth iterate. In other words, describe f^5 . What is g^3 if $g : \mathbb{R} \rightarrow \mathbb{R}$ is defined by the formula $g(x) = 2x^2 + 1$? Here \mathbb{R} is the set of real numbers (to be developed later in the course).

Exercise 14. Let $f : S \rightarrow S$ be given. Prove that f^2 is the identity function on S if and only if $f = f^{-1}$.

Note: Here we are using ‘ -1 ’ as a symbolic expression to mark the inverse function; it does not yet refer to a number. We will not define negative numbers until Chapter 4. Recall that f^{-1} is defined to be the inverse function of f , which exists if and only if f is bijective.

Exercise 15. Prove that $\sigma^3(2) = 5$ where $\sigma : \mathbb{N} \rightarrow \mathbb{N}$ is the successor function.

Informal Exercise 16. What is $\sigma^n(m)$?

Informal Exercise 17. Propose an informal definition of addition in terms of iteration of the successor function. Discuss how multiplication can be explained in terms of iteration of addition, and how exponentiation can be explained in terms of iteration of multiplication.

4. ADDITION

As mentioned above, we define addition in terms of iteration of successor. Informally, you get $m + n$ by starting with m and taking the successor n times. This idea motivates the formal definition.

Definition 8 (Addition). Let $m, n \in \mathbb{N}$. Let $\sigma^n : \mathbb{N} \rightarrow \mathbb{N}$ be the n th iteration of the successor map. Then

$$m + n \stackrel{\text{def}}{=} \sigma^n(m).$$

Observe that addition defines a function $\mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}$.

Remark 13. Functions $S \times S \rightarrow S$ are called *binary operations*. Thus $+$ is a binary operation on \mathbb{N} .

Remark 14. This is not the only way of viewing addition. In Chapter 3, we will show how $+$ can be understood in terms of counting the elements in a disjoint union.

The following are consequences of the iteration axiom and Definition 8.

Theorem 4. For all $m \in \mathbb{N}$

$$m + 0 = m.$$

Proof. By definition $m + 0 = \sigma^0(m)$. Recall that σ is a function $\mathbb{N} \rightarrow \mathbb{N}$. By the iteration axiom, σ^0 is the identity $i : \mathbb{N} \rightarrow \mathbb{N}$. Thus $\sigma^0(m) = m$ by definition of identity function. So $m + 0 = m$ by transitivity of equality. \square

Lemma 5. For all $m, n \in \mathbb{N}$

$$m + \sigma n = \sigma(m + n).$$

Exercise 18. Prove the above lemma.

Remark 15. As mentioned earlier a lemma is a kind of theorem whose purpose in life is to help prove more important theorems. The above result is relegated to the role of lemma not because it is not of independent interest, but because it will be superseded by a more general theorem (the associative law of addition), so its usefulness is only temporary.

Lemmas are not always simple. In fact, many times in mathematics a lemma will be more complicated to state or harder to prove than the main theorem. Part of the art of mathematics is to decide what lemmas to prove in order to make the proofs of the important theorems as clear and elegant as possible.

Remark 16. Many authors define addition in terms of recursion instead of iteration of successor. The above theorem and lemma are the two recursion conditions used in this approach.

Informally we know that successor σ is just addition by one. The following makes this official:

Theorem 6. For all $m \in \mathbb{N}$

$$m + 1 = \sigma m.$$

Exercise 19. Prove the above theorem.

Remark 17. From now we can replace σm with $m + 1$ whenever we want. Based on the above theorem, these two expressions are completely interchangeable.

Exercise 20. Use the above theorem to prove that $1 + 1 = 2$.

Exercise 21. Prove that $2 + 2 = 4$. Prove that $2 + 3 = 5$. Prove that $3 + 2 = 5$.

Now we come to the first major theorem of the chapter.

Theorem 7 (Associative Law). For all $x, y, z \in \mathbb{N}$

$$x + (y + z) = (x + y) + z.$$

Proof. Fix $x, y \in \mathbb{N}$, and let $S_{x,y} \subseteq \mathbb{N}$ be the set of $z \in \mathbb{N}$ with the property that $x + (y + z) = (x + y) + z$.

First we observe that $0 \in S_{x,y}$ since, by Theorem 4 (twice),

$$x + (y + 0) = x + y = (x + y) + 0.$$

Now assume $z \in S_{x,y}$. By Lemma 5 (several times) and our assumption,

$$\begin{aligned} x + (y + \sigma z) &= x + \sigma(y + z) \\ &= \sigma(x + (y + z)) \\ &= \sigma((x + y) + z) \\ &= (x + y) + \sigma z. \end{aligned}$$

So $\sigma z \in S_{x,y}$.

By the induction axiom, $S_{x,y} = \mathbb{N}$. This is true for any $x, y \in \mathbb{N}$. So if $x, y, z \in \mathbb{N}$ are arbitrary, $z \in S_{x,y}$ which implies $x + (y + z) = (x + y) + z$. \square

Remark 18. This proof by induction is valid, but, like many induction proofs, is weak on conveying an understanding *why* associativity is true. In Chapter 3 we give a second, more insightful proof involving the set theoretic identity $A \cup (B \cup C) = (A \cup B) \cup C$.

Warning: we do not yet have the commutative law. Thus the next two lemmas are not redundant. They do not merely repeat Theorems 4 and 6, but assert something truly new. They are lemmas since, once the commutative law is proved, they will become redundant. So they are only of temporary use.

Lemma 8. *If $n \in \mathbb{N}$ then $\sigma^n(0) = n$. In particular $0 + n = n$.*

Lemma 9. *If $n \in \mathbb{N}$ then $1 + n = \sigma n$.*

Proof. Let $S = \{x \in \mathbb{N} \mid 1 + x = \sigma x\}$. So $0 \in S$ since $1 + 0 = 1 = \sigma 0$.

Suppose $n \in S$.

$$\begin{aligned} 1 + \sigma n &= 1 + (n + 1) \\ &= (1 + n) + 1 \\ &= \sigma n + 1 \\ &= \sigma(\sigma n). \end{aligned}$$

So $\sigma n \in S$.

We conclude that $S = \mathbb{N}$. \square

Exercise 22. Prove Lemma 8. Complete the above sketchy proof of Lemma 9 by justifying every step by referring to earlier results, definitions, assumptions, or axioms, or by referring to the definition of S .

Theorem 10 (Commutative Law). *If $x, y \in \mathbb{N}$ then*

$$x + y = y + x.$$

Proof. Fix $x \in \mathbb{N}$, and let $S_x = \{u \in \mathbb{N} \mid x + u = u + x\}$. By Theorem 4 and Lemma 8, we get $0 \in S_x$.

Now assume $n \in S_x$. So

$$\begin{aligned} x + \sigma n &= \sigma(x + n) \\ &= \sigma(n + x) \\ &= 1 + (n + x) \\ &= (1 + n) + x \\ &= \sigma n + x \end{aligned}$$

We conclude that $\sigma n \in S_x$.

By the induction axiom, $S_x = \mathbb{N}$. This is true of any $x \in \mathbb{N}$ since x was chosen to be any arbitrary element of \mathbb{N} . Now let $x, y \in \mathbb{N}$ be any two elements of \mathbb{N} . Since $S_x = \mathbb{N}$, we have $y \in S_x$. By definition of S_x we conclude that $x + y = y + x$. \square

Remark 19. In Chapter 3 we see a more insightful proof of the commutative law involving the set theoretic identity $A \cup B = B \cup A$.

Exercise 23. Justify every step in the above proof by referring to earlier results, assumptions, or axioms, or by referring to the definition of S_x .

Exercise 24. Prove $(x + y) + z = (x + z) + y$ without using induction.

5. MULTIPLICATION

As mentioned above, our strategy for defining multiplication is to use iteration of addition. To understand, how this works, first consider the following familiar informal definition:

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

We can interpret the phrase “ n times” in terms of iteration. To see this, notice how we can build up to this sum in n steps:

STEP 1:	Add m to 0:	$0 + m$
STEP 2:	Add m to previous result:	$(0 + m) + m$
STEP 3:	Add m to previous result:	$((0 + m) + m) + m$
	\vdots	

Observe that we are just iterating the function $x \mapsto x + m$ as we go through the steps: every step involves applying $x \mapsto x + m$ where x is the result of the previous step. Observe also that the n th step results in $m \cdot n$, and that we start with $x = 0$ in the first step. (If we started with $x = m$, which might seem more natural, we would only use $n - 1$ steps. We prefer to take exactly n steps, so we want to start at $x = 0$). We call the function $x \mapsto x + m$ the “addition by m ” function or the “translation by m ” function, and we write it as α_m . Multiplication is obtained by iterating α_m . For the product $m \cdot n$, we iterate n times.

This informal discussion motivates the following formal definition:

Definition 9 (Multiplication). Let $m, n \in \mathbb{N}$. Let $\alpha_m : \mathbb{N} \rightarrow \mathbb{N}$ be defined by the rule $x \mapsto x + m$, and let α_m^n be the n th iteration of α_m . Then

$$m \cdot n \stackrel{\text{def}}{=} \alpha_m^n(0).$$

In particular, multiplication defines a binary operation $\mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}$. As is common, we do not always need to write the dot \cdot , but can use juxtaposition to indicate multiplication.

Remark 20. This is not the only way of viewing multiplication. In Chapter 3, we will show how multiplication can be thought of in terms of counting the elements in the Cartesian product of two finite sets. Another popular approach is through recursion (using the equations of Theorem 11 and Lemma 12).

Exercise 25. Prove the following theorem, lemma, and theorem using the iteration axiom and the definition of multiplication.

Theorem 11. For all $m \in \mathbb{N}$

$$m \cdot 0 = 0.$$

Lemma 12. For all $m, n \in \mathbb{N}$

$$m \cdot \sigma n = (m \cdot n) + m.$$

Theorem 13. For all $m \in \mathbb{N}$

$$m \cdot 1 = m.$$

Exercise 26. We already have proved that $2 + 2 = 4$ and $3 + 2 = 5$ (Exercise 21). Use the results concerning addition to give proofs of the following: $4 + 2 = 6$, $5 + 2 = 7$, $6 + 2 = 8$, $7 + 2 = 9$, $3 + 3 = 6$, $4 + 3 = 7$, $5 + 3 = 8$, $6 + 3 = 9$, $4 + 4 = 8$, $5 + 4 = 9$. Use these addition facts together with Theorems 11 and 13 and Lemma 12 to show the following: $0 \cdot 0 = 0$, $0 \cdot 1 = 0$, $0 \cdot 2 = 0$, $1 \cdot 1 = 1$, $1 \cdot 2 = 2$, $2 \cdot 0 = 0$, $2 \cdot 1 = 2$, $2 \cdot 2 = 4$, $2 \cdot 3 = 6$, $2 \cdot 4 = 8$, $3 \cdot 2 = 6$, $3 \cdot 3 = 9$.

Theorem 14 (Distributive Law: part 1). For all $x, y, z \in \mathbb{N}$

$$(x + y)z = xz + yz.$$

Remark 21. We adopt the usual conventions for dropping parentheses. Thus, when the parentheses and the dots are restored, the above equation is

$$(x + y) \cdot z = (x \cdot z) + (y \cdot z).$$

Exercise 27. Prove the distributive law. Do so by defining, for any fixed $x, y \in \mathbb{N}$, a set $S_{x,y} \subseteq \mathbb{N}$. Show that $S_{x,y} = \mathbb{N}$ by the axiom of induction. In order to give a complete and rigorous proof, *do not leave any parentheses out in this proof.*

Remark 22. This induction proof is valid but, like many induction proofs, weak on conveying an understanding *why* the result is true. In Chapter 3 we will see a second proof using the set theoretic identity

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

Lemma 15. *If $n \in \mathbb{N}$ then $0 \cdot n = 0$.*

Lemma 16. *If $n \in \mathbb{N}$ then $1 \cdot n = n$.*

Exercise 28. Prove the above two lemmas using the induction axiom.

Theorem 17 (Commutative Law). *For all $x, y \in \mathbb{N}$*

$$xy = yx.$$

Proof. Fix $x \in \mathbb{N}$. Let $S_x = \{u \in \mathbb{N} \mid xu = ux\}$. We wish to show $y \in S_x$. We do so by showing all natural numbers are in S_x (via induction).

By Theorem 11 and Lemma 15, we get $0 \in S_x$.

Now assume $n \in S_x$. Then

$$\begin{aligned} x \cdot \sigma n &= xn + x \\ &= nx + x \\ &= n \cdot x + 1 \cdot x \\ &= (n + 1) \cdot x \\ &= \sigma n \cdot x. \end{aligned}$$

We conclude that $\sigma n \in S_x$.

By the induction axiom, $S_x = \mathbb{N}$. Thus $y \in S_x$ which implies $xy = yx$. \square

Exercise 29. Justify every step in the above proof.

Remark 23. In Chapter 3 we give a more insightful proof involving the natural bijection from $A \times B$ to $B \times A$.

Corollary 18 (Distributive Law: part 2). *For all $x, y, z \in \mathbb{N}$*

$$x(y + z) = xy + xz.$$

Exercise 30. Prove the above corollary using the commutative law, and without using induction.

Exercise 31. Try to prove the following without looking at the given proof. If you get stuck, take a short peek at the proof for ideas. Now compare your proof to the given proof. Justify every step in the given proof.

Theorem 19 (Associative Law). *For all $x, y, z \in \mathbb{N}$*

$$x(yz) = (xy)z.$$

Proof. Let $S_{x,y} = \{u \in \mathbb{N} \mid x(yu) = (xy)u\}$.

First we check that $0 \in S_{x,y}$. This follows from Theorem 11.

Now assume $n \in S_{x,y}$. So

$$\begin{aligned} x(y(n+1)) &= x((y \cdot n) + (y \cdot 1)) \\ &= x((yn) + y) \\ &= (x(yn)) + (xy) \\ &= ((xy)n) + ((xy)1) \\ &= (xy)(n+1) \end{aligned}$$

So $\sigma n = n + 1$ is in $S_{x,y}$.

By the induction axiom, $S_{x,y} = \mathbb{N}$. In particular, $z \in S_{x,y}$. \square

6. EXPONENTIATION

Just as repeated addition gives multiplication, repeated multiplication gives exponentiation. In other words, you can define exponentiation via the iteration of a multiplication function. How we do this for exponentiation is similar to how we developed multiplication, so the details will be left to the reader.

Definition 10. Let $m, n \in \mathbb{N}$. Let $\mu_m : \mathbb{N} \rightarrow \mathbb{N}$ be defined by the rule $x \mapsto xm$. Let μ_m^n be the n th iteration of μ_m . Then

$$m^n \stackrel{\text{def}}{=} \mu_m^n(1).$$

Remark 24. One amusing aspect of our approach is that exponential notation is used for iteration (Section 3) before it is used in the traditional way for exponentiation itself (here in Section 6). This is a symptom of the large emphasis we place on functions and their iterates. Our convention is that when an exponent is used with a function it refers to iteration, but when it is used with a number it refers to exponentiation.

Informal Exercise 32. In contrast with the previous section, we start with 1 instead of 0 in our iterative definition. What would happen if we used 0 instead of 1 in Definition 10?

Remark 25. This is not the only way of viewing exponentiation. In Chapter 3, we will see how it can be defined in terms of counting the number of functions between two sets. It can also be defined using recursion.

Informal Exercise 33. Do you expect $(m, n) \mapsto m^n$ to be a commutative binary operation $\mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}$? Do you expect it to be associative? If you said ‘no’ to either question, back up your answer with a counter-example.

Theorem 20. For all $m \in \mathbb{N}$

$$m^0 = 1.$$

Lemma 21. For all $m, n \in \mathbb{N}$

$$m^{\sigma n} = m^n \cdot m$$

Theorem 22. For all $m \in \mathbb{N}$

$$m^1 = m.$$

Exercise 34. Prove the above theorems and lemmas.

Exercise 35. Use the above theorems and lemmas to show the following:
 $0^0 = 1$, $2^2 = 4$, $2^3 = 8$.

Warning. Although the equation $0^0 = 1$ is valid in our current context, there are some parts of mathematics where 0^0 is regarded as undefined. This is related to the use of limits in calculus where we have to be careful with limits that converge to indeterminate expressions of the form $0/0$, ∞/∞ , $\infty - \infty$, or even 0^0 . Limits of expressions in indeterminate form do not consistently converge to any fixed value. In fact, some limits in indeterminate form diverge, and some converge, and those that converge do not all converge to the same value. The problem with limits with indeterminate form 0^0 is related to the fact that the function $f(x, y) = x^y$ is not continuous at $(0, 0)$. So in calculus and other contexts, 0^0 is often left undefined.

Theorem 23. If $x, y, n \in \mathbb{N}$ then

$$(xy)^n = x^n y^n.$$

Exercise 36. Prove the above using induction on n . In other words, apply the induction axiom to a certain set $S_{x,y}$.

Theorem 24. If $x, m, n \in \mathbb{N}$ then

$$x^{m+n} = x^m x^n.$$

Exercise 37. Prove the above using induction on n . In other words, apply the induction axiom to a certain set $S_{x,m}$.

Theorem 25. If $n \in \mathbb{N}$ is not 0 then

$$0^n = 0.$$

Exercise 38. Prove the above *without* induction. Use Theorem 2 to first show that $n = \sigma m$ for some m .

Theorem 26. If $n \in \mathbb{N}$ then

$$1^n = 1.$$

Theorem 27. If $x, n, m \in \mathbb{N}$ then

$$(x^m)^n = x^{mn}.$$

Exercise 39. Prove the above two theorems.

7. OTHER PROPERTIES OF ADDITION

There are a few more facts about addition that we will need in the next chapter.¹²

Theorem 28. *If $m, n \in \mathbb{N}$ are such that $0 = m + n$ then $m = n = 0$.*

Proof. Suppose $n \neq 0$. Then $n = \sigma(x)$ for some $x \in \mathbb{N}$. Thus

$$0 = m + n = m + \sigma(x) = \sigma(x + m).$$

This contradicts Axiom 2. From this contradiction, we conclude $n = 0$.

From $n = 0$ and $0 = m + n$ we get $m = 0$ as well using Theorem 4. \square

Exercise 40. Supply the missing justifications in the above proof, including every step in the chain of equalities

$$0 = m + n = m + \sigma(x) = \sigma(x + m).$$

By definition of addition, we know that \mathbb{N} is closed under addition. In other words, if $a, b \in \mathbb{N}$ then $a + b \in \mathbb{N}$. The following shows that \mathbb{N}^+ is closed as well. Recall that \mathbb{N}^+ is the set of nonzero natural numbers (Definition 4).

Corollary 29. *The set \mathbb{N}^+ of positive natural numbers is closed under addition. In other words, if $m, n \in \mathbb{N}^+$ then $m + n \in \mathbb{N}^+$.*

Exercise 41. Prove the above corollary.

Now we turn our attention to the cancellation law for addition. Both addition and multiplication have cancellation laws. For addition the law states that if $x + z = y + z$ then $x = y$. In other words, we cancel z . Contrary to what one might think, the operation of subtraction is not needed to state or prove this law. For multiplication the law states that if $xz = yz$ and $z \neq 0$ then $x = y$. Note the extra condition $z \neq 0$. The cancellation law for multiplication will be proved in the next chapter.

First we consider the contrapositive form (which is a bit easier to prove):

Theorem 30. *Suppose $x, y \in \mathbb{N}$ are distinct: $x \neq y$. Then $x + z \neq y + z$ for all $z \in \mathbb{N}$.*

Proof. Fix $x, y \in \mathbb{N}$ distinct, and let $S_{x,y} = \{z \in \mathbb{N} \mid x + z \neq y + z\}$. Since $x + 0 = x$ and $y + 0 = y$, we have $x + 0 \neq y + 0$. So $0 \in S_{x,y}$.

Suppose that $n \in S_{x,y}$. We wish to show that $\sigma n = n + 1$ is in $S_{x,y}$, i.e., that $x + \sigma(n) \neq y + \sigma(n)$. Since $n \in S_{x,y}$ we have $x + n \neq y + n$. Since σ is injective, we have $\sigma(x + n) \neq \sigma(y + n)$. Observe that by Lemma 5

$$\sigma(x + n) = x + \sigma(n), \text{ and}$$

$$\sigma(y + n) = y + \sigma(n).$$

Thus $x + \sigma(n) \neq y + \sigma(n)$. In other words, $\sigma(n) \in S_{x,y}$.

By the induction axiom, $S_{x,y} = \mathbb{N}$. Thus, $x + z \neq y + z$ for all $z \in \mathbb{N}$. \square

¹²Nothing in this section uses the results of Section 5 or 6.

Theorem 31 (Cancellation Law). *Suppose $x, y, z \in \mathbb{N}$. Then*

$$x + z = y + z \quad \text{implies} \quad x = y.$$

Proof. Suppose that $x + z = y + z$, but that $x \neq y$. By Theorem 30, this implies that $x + z \neq y + z$, a contradiction. \square

In Exercise 11 we saw that $n \neq n + 1$. The following generalizes this to other sums. It can be used to show, for instance, that the natural numbers we have defined so far, $0, 1, 2, 3, 4, \dots, 9$, are pairwise distinct.

Theorem 32. *Suppose $n, m, b \in \mathbb{N}$. If $n = m + b$ where $b \neq 0$ then $n \neq m$.*

Proof. Suppose $n = m + b$ where $b \neq 0$, but that $n = m$. So

$$b + m = m + b = n = m = 0 + m.$$

By the cancellation law $b = 0$, contradicting our hypothesis. \square

8. THE UNIVERSAL PROPERTY (OPTIONAL)

Earlier we adopted the iteration axiom and used it to help define and prove the basic properties of addition, multiplication, and such. However, a promise was made to show that the iteration axiom is not needed since it can be derived from the Peano axioms. In this section we prepare for a proof of the iteration axiom by making a careful study of iteration.

In order to avoid circularity we will appeal only to the Peano axioms, and not to any theorems proved with the assistance of the iteration axiom. In fact, this section and the next could be cut and pasted immediately after Section 2 with no loss of logical rigor. I did not do this since I felt that the development would go more smoothly if we *applied* iteration to define and prove things about addition, multiplication, and such before giving the more elaborate proofs justifying iteration.

Let's begin with an informal discussion about iteration. The raw materials of iteration consists of a set A and a function $s : A \rightarrow A$. If you choose a starting element $z \in A$ and repeatedly apply s you will get

$$z, \quad s(z), \quad s(s(z)), \quad s(s(s(z))), \quad s(s(s(s(z))))),$$

and so on. We can informally think of this as a sort of "path", and we can think of s as determining a "step". Metaphorically, you are starting with z and stepping along a path away from z .

An example of this is the definition of $m + n$ by iteration of $\sigma : \mathbb{N} \rightarrow \mathbb{N}$. In that case we start with $z = m$ and iterate σ a total of n times. In other words, we take n steps along the path to reach our goal of $m + n$. Here each step consists of applying σ to the previous result.

Fix $z \in A$ and $s : A \rightarrow A$ as above. The following theorem defines a function $\varphi : \mathbb{N} \rightarrow A$ that in some sense identifies the result of taking n steps starting from z . In other words,

$$\varphi(0) = z, \quad \varphi(1) = s(z), \quad \varphi(2) = s(s(z)), \quad \varphi(3) = s(s(s(z))),$$

and so on. This function will depend, of course, on z and s , and we could write it as $\varphi_{z,s}$ if we want to make this dependency clear. It turns out that the condition $\varphi \circ \sigma = s \circ \varphi$ is what is needed to force φ to be such a “stepping function” (see the corollary).

This informal discussion helps motivate the following theorem. It states that it is possible to find such a function φ .

Theorem 33 (Universal Property of \mathbb{N}). *Suppose A is a set, $z \in A$ is an element, and $s : A \rightarrow A$ is a function. Then there is a unique function $\varphi : \mathbb{N} \rightarrow A$ such that $\varphi(0) = z$ and $\varphi \circ \sigma = s \circ \varphi$.*

Proof. Let $\sigma_2 : \mathbb{N} \times A \rightarrow \mathbb{N} \times A$ be defined by the rule $(n, a) \mapsto (\sigma n, sa)$. (It is called σ_2 since it takes pairs to pairs). For a subset R of $\mathbb{N} \times A$, we say R is z -closed if (i) $(0, z) \in R$ and (ii) $\sigma_2[R] \subseteq R$. This second condition means that if $(n, a) \in R$ then $(\sigma n, sa)$ must be in R .

Observe also that the intersection of z -closed sets is z -closed. Observe that whole set $\mathbb{N} \times A$ is z -closed. Let N be the intersection of all z -closed sets (the z -closure), so N is itself closed. Since it is the intersection of all z -closed sets, N is contained in any given z -closed set.

Since $(0, z) \in N$ and $\sigma_2[N] \subseteq N$, it follows that $\sigma_2[N] \cup \{(0, z)\} \subseteq N$. Check that $\sigma_2[N] \cup \{(0, z)\}$ is z -closed. So, by the minimality of N ,

$$N = \sigma_2[N] \cup \{(0, z)\}.$$

We say $n \in \mathbb{N}$ is *paired-up* if N has a unique pair (n, a) with first coordinate n . In this case, we say that a *corresponds to* n . Claim: all $n \in \mathbb{N}$ are paired up. We prove this claim by induction.

Since $N = \sigma_2[N] \cup \{(0, z)\}$, it follows that 0 is paired-up: Axiom 2 implies that no pair of the form $(0, a)$ is in the image of σ_2 . Also, z corresponds to 0.

Suppose that n is paired-up. Then there is a unique pair $(n, a) \in N$. Thus $(\sigma n, sa)$ is in N since N is z -closed. We now want to show $(\sigma n, sa)$ is the unique pair with first coordinate σn . So, suppose $(\sigma n, b)$ is also in N . Since $N = \sigma_2[N] \cup \{(0, z)\}$ and since $\sigma n \neq 0$ it follows that $(\sigma n, b) \in \sigma_2[N]$. So $(\sigma n, b) = (\sigma m, sc)$ for some pair $(m, c) \in N$. Since σ is injective, $m = n$. So $(n, c) \in N$. Observe that (n, a) and (n, c) are in N . But n is paired-up, so, by uniqueness, $c = a$. Thus $(\sigma n, b) = (\sigma m, sc) = (\sigma n, sa)$. This concludes the argument for uniqueness and shows that σn is paired-up.

Let $S \subseteq \mathbb{N}$ be the subset of paired-up elements. By the induction axiom, $S = \mathbb{N}$. Thus every natural number is paired-up.

Let φ be defined by the rule $n \mapsto a$ where a corresponds to n . In other words φn corresponds to n , so $(n, \varphi n) \in N$. Since N is z -closed $(\sigma n, s(\varphi n)) \in N$. Thus $s(\varphi n)$ corresponds to σn . So $\varphi(\sigma n) = s(\varphi n)$ by the definition of φ . This holds for all $n \in \mathbb{N}$, so we have established that $\varphi \circ \sigma = s \circ \varphi$. Since z corresponds to 0, we have $\varphi(0) = z$. We have now established the existence of the desired φ .

We now show uniqueness. Suppose that $\varphi' : \mathbb{N} \rightarrow A$ is such that $\varphi'(0) = z$ and $\varphi' \circ \sigma = s \circ \varphi'$. We need to show that $\varphi = \varphi'$. If W is the set of $n \in \mathbb{N}$ such that $\varphi(n) = \varphi'(n)$, we need to show $W = \mathbb{N}$. We do this by induction.

Observe that $\varphi(0) = z = \varphi'(0)$, so $0 \in W$. Now assume that $n \in W$, so $\varphi(n) = \varphi'(n)$. Then $s(\varphi(n)) = s(\varphi'(n))$. However,

$$\varphi(\sigma n) = (\varphi \circ \sigma)(n) = (s \circ \varphi)(n) = s(\varphi(n)),$$

and

$$\varphi'(\sigma n) = (\varphi' \circ \sigma)(n) = (s \circ \varphi')(n) = s(\varphi'(n)).$$

So $\varphi(\sigma n) = \varphi'(\sigma n)$. In particular, $\sigma n \in W$.

By the induction axiom, $W = \mathbb{N}$. So, $\varphi = \varphi'$. \square

Exercise 42. Let A, s, z, φ be as in the above theorem. Show the following

$$\varphi(0) = z, \quad \varphi(1) = s(z), \quad \varphi(2) = s(s(z)), \quad \varphi(3) = s(s(s(z))).$$

Informal Exercise 43. Describe φ in the case where $A = \{0, 1\}$, $z = 0$, and $s : A \rightarrow A$ is defined by the rule $0 \mapsto 1$ and $1 \mapsto 0$. What if $z = 1$ instead? What if s is the identity function instead?

Remark 26. This theorem is called the *universal* property of \mathbb{N} . To explain this, we need to discuss some ideas related to category theory.

In the last 60 years or so, mathematicians have become more concerned with the notion of *structure*. Roughly speaking, a structure is typically a set equipped with special relations, functions, binary operators, elements, and the like. The field of mathematics that is used to compare structures is *category theory*. We will not discuss category theory in general, but will illustrate some ideas of category theory in the context of the above theorem.

Recall that the first of the Peano axioms describes \mathbb{N} as a set equipped with two things (i) a starting element $0 \in \mathbb{N}$ and (ii) a function (called the successor function) $\sigma : \mathbb{N} \rightarrow \mathbb{N}$, that can be used to “take steps” away from the starting point 0. Similarly, the set A in the theorem is given with a function $s : A \rightarrow A$ and a starting element z . We can view \mathbb{N} and such a set A as examples of a certain basic type of structure. Let’s make up some fancy terminology and call such a structure a *path structure* since we indicate a starting point and from there can go on a path through the set by using the function to take steps. More precisely, a *path structure* is a set A equipped with (i) a *starting point* $z \in A$ and (ii) a *stepping function* $s : A \rightarrow A$.

A simple example of a path structure is the set $A = \{0, 1\}$ where we declare the starting point to be 0 and declare the stepping function to be the function $s : A \rightarrow A$ defined by the rule $0 \mapsto 1$ and $1 \mapsto 0$. Repeating s gives you the path $0, 1, 0, 1, 0, \dots$. As you can see, we do not require that every path structure satisfy all the Peano axioms. For example, in this structure 0 is in the image of the stepping function.

We use different path structures for different situations. The path structure most appropriate for defining m^n is that given by taking the set \mathbb{N} ,

but declaring the starting point to be 1, and declaring the stepping function to be $\mu_m(x) = xm$. Following a path in this structure would give you $1, m, m^2, m^3, m^4, \dots$

The collection of all possible path structures forms something that mathematicians call a *category*.

Now, among all path structures, \mathbb{N} is very special: Theorem 33 shows it maps (uniquely) to any other path structure in a special way. More specifically, given any other path structure A , there is a function $\varphi : \mathbb{N} \rightarrow A$ such that (i) $0 \mapsto z$ and (ii) $\varphi \circ \sigma = s \circ \varphi$. The first condition says that φ sends the start to the start. The second condition matches σ with s . We can illustrate the second condition with the following *commutative diagram*:

$$\begin{array}{ccc} \mathbb{N} & \xrightarrow{\varphi} & A \\ \downarrow \sigma & & \downarrow s \\ \mathbb{N} & \xrightarrow{\varphi} & A \end{array}$$

What this diagram expresses is that both ways of going from the top left set to the bottom right set gives the same image.¹³ This diagram expresses the equation $\varphi \circ \sigma = s \circ \varphi$.

The existence of this special φ is called the *universal property of \mathbb{N}* . In other words, \mathbb{N} has the universal property of being able to map uniquely into any other path structure (in a path compatible way).

Remark 27. The universal property (Theorem 33) is important for other reasons besides describing iteration. In fact, it makes it easy to show that any two models of the Peano axioms are “isomorphic”. However, we will skip this important isomorphism theorem since explaining in precisely will lead us too far afield.

9. ELIMINATING THE ITERATION AXIOM (OPTIONAL)

We now use Theorem 33 to show that the iteration axiom can be dispensed with. In other words, it can be proved as a theorem.

Theorem 34. *Let $f : S \rightarrow S$ be a function. Then one can assign to every $n \in \mathbb{N}$ a function $f^n : S \rightarrow S$ such that (i) f^0 is the identity function on S , and (ii) $f^{\sigma n} = f \circ f^n$.*

The idea behind this theorem is to use Theorem 33 to describe iteration. If $a \in S$ then we want to consider the iteration process giving

$$a, \quad f(a), \quad f(f(a)), \quad f(f(f(a))), \quad f(f(f(f(a)))),$$

and so on. So we apply the theorem to the case where $A = S$, $z = a$, and $s = f$. Then $f^n(a)$ is obviously $\varphi(n)$. We give the details below:

¹³In category theory the map φ is called a *morphism* or a *homomorphism* because it in some sense preserves the form (“morph”) of the structures. Different categories have different types of morphisms. For example, in the category of vector spaces, the morphisms are linear maps: they preserve the linear structure.

Proof. Define f^n by the rule $a \mapsto \varphi_a(n)$ where φ_a is the function φ described in Theorem 33 given by choosing $A = S$, $s = f$, and $z = a$. Observe that f^n sends elements $a \in S$ to elements of S .

Observe also that $f^0(a) = \varphi_a(0) = a$ (because $z = a$ in this case). Thus f^0 is the identity function.

Finally, use Theorem 33 to observe that

$$f^{\sigma n}(a) = \varphi_a(\sigma n) = f(\varphi_a(n)) = f(f^n(a)) = f \circ f^n(a)$$

(recall $s = f$). This holds for arbitrary $a \in S$, so $f^{\sigma n} = f \circ f^n$. \square

There is another proof that is interesting (also based on Theorem 33, but with different choice of A, z and $s : A \rightarrow A$):

Proof. (Second proof) Let A be the set of functions from S to itself. Let z be the identity function on S . Let $s : A \rightarrow A$ be the map that sends a function g to $f \circ g$.

Let $\varphi : \mathbb{N} \rightarrow A$ be as in Theorem 33. Define f^n to be $\varphi(n)$. Since A consists of functions from S to itself, f^n maps S to S . Since $\varphi(0) = z$, we have that f^0 is the identity function on S . Finally, since $\varphi \circ \sigma = s \circ \varphi$,

$$f^{\sigma n} = \varphi(\sigma n) = \varphi(\sigma(n)) = s(\varphi(n)) = s(f^n) = f \circ f^n.$$

This completes the proof. \square

Remark 28. There is also a uniqueness result. Let $\text{maps}(S, S)$ be the set of all functions $S \rightarrow S$ (written A in the second proof). Then the theorem describes the existence of a function $\mathbb{N} \rightarrow \text{maps}(S, S)$, given by $n \mapsto f^n$, that satisfies certain properties. The second proof above can be modified and extended to show the uniqueness of the function $\mathbb{N} \rightarrow \text{maps}(S, S)$ with the desired properties.