Math 3040 **Spring 2009**

The Natural Numbers[†]

Peter J. Kahn

[Revised: Mon, Jan 26, 2009]

Contents

1. History 2
2. The basic construction 2
3. Proof by Induction 7
4. Definition by induction 12
5. Addition 14
6. Addition and the natural ordering 19
7. Variants on induction 20
8. Multiplication 23

Proofs are used both to assist or insure the solution of specific problems in mathematics and to provide the cement that holds together the construction of large-scale mathematical structures, such as the theory of real numbers or the theory of groups.

In this and subsequent chapters, we shall be using proofs for a large-scale project, namely the construction of some of the most common number systems used in mathematics. We shall begin by constructing the natural numbers and then follow a classical path through the construction of the integers and rationals. The path then branches, with one branch leading to the development of the real numbers and the beginnings of analysis and the other with an algebraic focus. In part, this project may serve as an introduction to some important mathematical ideas. But in equal measure, it is intended as a display of proof techniques and of how proofs are used to build mathematical theories.

We do all this from scratch using the rules of mathematical logic, as described earlier, together with as few of the basic concepts of intuitive set theory as possible. Our intention is to give a relatively complete and rigorous development appropriate for the level of this course. This means that numerous subtle, foundational aspects will not get as thorough a treatment as they might deserve and would certainly receive in a more advanced course in mathematical logic or set theory.

1. History

So-called *counting numbers* have been used, with varied symbolic representation, since prehistoric times. Notation for these has stabilized to the standard Arabic numerals $1, 2, 3, \ldots$ for the past millennium or so, and classical usage has referred to these numbers as "natural." The concept of zero achieved acceptance as a number much more recently—in roughly 300 B.C.E.—and its common denotation as 0 occurred well after the early usage of $1, 2, 3 \ldots$. So, it may be deemed to be slightly less natural than the others. Nevertheless, we (and most contemporary mathematicians) adjoin 0 to the counting numbers, call the totality $\{0, 1, 2, 3, \ldots\}$ the *set of natural numbers*, and denote it by the letter $\mathbb{N}$.

The foregoing refers to a *chronological* or historical development, as opposed to the *logical* development that is the goal of this chapter. Of course we shall use our understanding of the natural numbers handed down to us through history to guide and illustrate the logical construction we present.

2. The basic construction

There are a number of different approaches to constructing the natural numbers. These are all essentially equivalent. They vary according to what one considers to be an appropriate starting point and what one allows as reasonable rules of formation.

The approach taken here is to construct the natural numbers within a universe $\mathcal{U}$ about which we assume as little as possible. In *Set Theory* we said simply that $\mathcal{U}$ is a set whose members consist of all the objects that could possibly interest us. For our purposes here, it will be sufficient to assume simply that $\mathcal{U}$ is a non-vacuous family of sets that is closed under the operations of set formation, subset, binary cartesian product, and power set. We also want $\mathcal{U}$ to be closed under the operation of indexed union (as described in *Set Theory*), where the index set can be any set in $\mathcal{U}$. We make no further assumption now about the nature of the sets that are members of $\mathcal{U}$, but we take note of the fact that since the empty set $\emptyset$ is a subset of every set, we do know that $\emptyset$ belongs to $\mathcal{U}$. In fact, at the outset, this is the only specific set in $\mathcal{U}$ whose identity we know. Of course, the existence of $\emptyset$ as a member of $\mathcal{U}$ entails that all sets we can obtain from $\emptyset$ by finite applications of the above-described constructions must also belong to $\mathcal{U}$.

Now let us be more specific and begin our construction of the natural numbers. Since we are using sets in this construction, we must find some sets in $\mathcal{U}$ that are natural candidates for the usual numbers. Clearly, the empty set is a natural candidate for representing what we want to think of as 0, and fortunately, as we observed, $\emptyset$ is a member of $\mathcal{U}$. So, we begin with that. Then we use the rules of formation mentioned and form sets described in Exercise 1 of *Set Theory*. That is, in addition to $\emptyset$, we get $\{\emptyset\}$, and $\{\emptyset, \{\emptyset\}\}$, both also in $\mathcal{U}$. For one more example, the reader should be able to see easily that the set $\{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$ can be formed in this way and that it is distinct from each of the other three.

Exercise 1. Prove that $\{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$ is distinct from the first three sets listed.

Notice that these sets contain $0, 1, 2, 3$ elements, respectively. Since these sets are in $\mathcal{U}$, they may be used as index sets for the union operation, as can any other set we

construct in this way. In particular, for example, the union of any two or three sets in $\mathcal{U}$ is again in $\mathcal{U}$.

We now continue with this sequential, constructive process using at each step the same formation rule that we used for the sets already constructed: that is, having constructed the finite set S as a member of $\mathcal{U}$, the next set is defined to be the union $S \cup \{S\}$.

Since $\mathcal{U}$ is closed under the operation of set formation and indexed union, this new set is also in $\mathcal{U}$. Moreover, by Proposition 2 of *Set Theory* , $S \cup \{S\}$ is also finite. We call it *the successor* of S and denote it by $\sigma(S)$.

successor function

To summarize,

$$\sigma(S) = S \cup \{S\}.$$

We may continue this process as long as we like, *but our assumptions so far do not allow us to conclude that the totality of all sets constructed by this sequential process is a set in* $\mathcal{U}$. We need an additional principle which allows this, an *et cetera principle* so to speak.

The Natural Number Axiom: All the sets constructed by the above-described sequential process, and only those, comprise a set in $\mathcal{U}$ that we denote by $\mathbb{N}$ and call the set of natural numbers. Each member of $\mathbb{N}$, that is, each of the constructed sets, is called a *natural number*.

This axiom has some simple consequences that allow us to identify this so-far very abstract definition of $\mathbb{N}$ with the more familiar concept that we have of the natural numbers. We begin with some simple observations. First, every collection of sets is partially ordered by the relation of set-inclusion. So, this relation defines a partial ordering on $\mathbb{N}$. But because of the way in which $\mathbb{N}$ is constructed, we can conclude

more. Each non-empty set T in our construction is a finite set which is obtained from any earlier constructed set S by adjoining some elements, i.e. $S \subseteq T$. Therefore, the partial

order on $\mathbb{N}$ is a linear order, because, given any two constructed sets S and T, either they are equal or one was constructed before the other. In either case, one is included in the other. We call this the *natural linear order on* $\mathbb{N}$.

natural linear order

However, even if sets are constructed at different points in our sequential procedure, we are not yet able to conclude that they are different. To verify that they are different, we need the following lemmas:

Lemma 1. *Let S and T be two of the constructed sets, with S constructed before T. Then $S \in T$.*

Proof. If T is constructed immediately after S, then $T = \sigma(S)$, and $S \in T$, by definition of $\sigma(S)$. If T is constructed later, then $\sigma(S) \subseteq T$. Since $S \in \sigma(S)$, it follows that $S \in T$. □

Lemma 2. *Let T be one of the constructed sets. Then T consists precisely of all the sets constructed earlier.*

Proof. Let B be the collection of all sets constructed before T. By definition of the construction, T is formed by adjoining to $\emptyset$ some or all of the members of B. So, $T \subseteq B$. On the other hand, Lemma 1 immediately implies that $B \subseteq T$. So, $T = B$. □

Lemma 3. *None of the constructed sets S satisfies $S \in S$.*

Proof. Each constructed S makes its *first* appearance at some definite step in our process. If this is at the first step, then $S = \emptyset$ and the result holds since then S has no members. If this is at a later step, then the members of S are sets that were constructed earlier. Since S is not one of these, $S \notin S$. □

Remark: Lemmas 1 and 3 imply that for any constructed sets S and T, with S constructed before T, we have $S \subset T$, i.e., we have strict inclusion. We also call this strict ordering the natural ordering on $\mathbb{N}$.

Exercise 2. Suppose that S and T are two constructed sets with S constructed earlier than T. Prove that $S \cup \{S\} \neq T \cup \{T\}$. (Hint: Use proof by contradiction. That is, assume that equality holds. Then, since $T \neq S$, conclude that $T \in S$. Show that this is impossible.)

Exercise 3. The successor rule $\sigma(S) = S \cup \{S\}$ defines a function from $\mathbb{N}$ to $\mathbb{N}$ which we denote by $\sigma : \mathbb{N} \to \mathbb{N}$. (a) Prove that σ is injective but not bijective. This proves that $\mathbb{N}$ is infinite, according to our definition in §4 of *Set Theory*. (b) Prove that if $S \subset T$, then $\sigma(S) \subset \sigma(T)$. In other words, $\sigma : \mathbb{N} \to \mathbb{N}$ preserves the natural ordering.

Proposition 1. *Let M be a non-empty subset of $\mathbb{N}$. Then M has a smallest member.*

Proof. This proof is based on our assumption about the sequential construction process: namely that each natural number in the sequence is constructed at a definite, finite step in the process. Let T be a member of the non-empty set M. T is a natural number, so it was constructed at a unique step in the process described, and it is a finite set whose members are all the natural numbers constructed earlier. If some of these belong to M, then clearly the one that was constructed first will be the smallest member of M. This natural number can be identified by a direct, finite inspection. If there are no such members, then obviously T is the smallest member of M. This completes the proof. □

Remark: Any linear ordering having the property described in Proposition 1 is called a *well-ordering*. Thus, Proposition 1 may be phrased as: The natural ordering on the

natural numbers is a well-ordering. Proposition 1 is also sometimes referred to as the **Well-Ordering Principle for the Natural Numbers**.

Proposition 2. ***(Principle of Induction:)*** *Suppose that X is a subset of $\mathbb{N}$ with the following two properties: (a) $\emptyset \in X$; (b) if $S \in X$, then $\sigma(X) \in X$.*
Then $X = \mathbb{N}$.

principle of induction

Proof. We assume hypotheses (a) and (b). Let Y be the complement of X in $\mathbb{N}$. We first give a proof by contradiction that Y is empty.

If Y is non-empty, then by Proposition 1, Y has a smallest member, say T, which occurs at some well-defined step in the construction of $\mathbb{N}$. It cannot occur at the initial step, because then it would equal $\emptyset$, which is assumed to belong to X. Therefore, it equals the successor $\sigma(S)$ of some set $S \in \mathbb{N}$. This S, being strictly smaller than T, cannot be in Y, so $S \in X$. But assumption (b) then implies that $\sigma(S) \in X$, i.e., $T \in X$, contradicting $T \in Y$. This contradiction proves that Y is empty.

But then $X = \mathbb{N} \setminus Y = \mathbb{N} \setminus \emptyset = \mathbb{N}$, as desired. □

3. Proof by Induction

It is now worth pausing in our construction to introduce the standard notation and terminology and to work a bit with the Principle of Induction.

It is standard to denote the first ten sets constructed by the procedure described above by the numerals $0, 1, 2, 3, 4, 5, 6, 7, 8, 9$, and then to form additional numerals via the well-known positional notation. Subsequently constructed sets are then denoted by these in order. So, we now assume this is done. We shall refer to these in the usual way as particular natural numbers. General elements of $\mathbb{N}$ will often be denoted by lower case letters such as $a, b, c, \ldots, i, j, k, l, m, n$, etc. We still use the notation σ for the successor function, but we'll often find it convenient to abbreviate $\sigma(n)$ as n'.

Since the natural numbers are each defined to be specific sets, the reader might be interested to know what these sets look like in standard notation. Of course 0 is just the empty set $\emptyset$. However, each successor number n' can be nicely written as

$$n' = \{0, 1, 2, \ldots, n\}.$$

Thus, $1 = \{0\}$, $2 = \{0, 1\}$, $3 = \{0, 1, 2\}$, etc.

We shall rarely if ever use the fact that the natural numbers are sets or use this representation of natural numbers as sets.

linear ordering of $\mathbb{N}$ restated

The natural well-ordering on $\mathbb{N}$ will be denoted by $\leq$ (and the strict ordering corresponding to this by $<$). Exercise 5 (b) may be rephrased as follows in our new notation: for all m and n in $\mathbb{N}$,

$$m < n \implies m' < n'.$$

This fact will be used later.

We have now recovered all of the notation that we usually use in connection with the natural numbers, short of that which relates to arithmetic.

Please note that we are still in a pre-arithmetic mode. This means that we may not yet make use of anything relating to addition, subtraction, multiplication, or division for our logical development. (Of course, we may well use arithmetic in illustrative examples or in exercises.)

Now let us reformulate the Principle of Induction, using the notation just introduced.

Principle of Induction for Sets: Let X be a subset of $\mathbb{N}$ satisfying the following two properties: (a) $0 \in X$; (b) $n \in X \implies n' \in X$. Then $X = \mathbb{N}$.

This version of the Principle of Induction is closer to what the student has already seen, and it is sometimes a useful formulation. However, this is still not the familiar version. To tie the Principle of Induction for Sets to the more familiar version, we recall the notion of the *truth set* of a predicate.

Let $P(n)$ be a predicate involving a single variable n that ranges over $\mathbb{N}$. The truth set of $P(n)$ is defined to be the set of all $n \in \mathbb{N}$ for which $P(n)$ is true. To prove that $P(n)$ is true for all n amounts to showing that the truth set of $P(n)$ equals $\mathbb{N}$. By the Principle of Induction for Sets this amounts to showing that the truth set satisfies (a) and (b) above. But each of (a) and (b) can be rephrased directly in terms of $P(n)$, thus producing the standard version of the Principle of Induction.

The Standard Principle of Induction: Let $P(n)$ be any predicate involving a free single variable n that ranges over $\mathbb{N}$. Then $(\forall n)P(n)$ is true if and only if (a) $P(0)$ is true and (b)$(\forall k)(P(k) \implies P(k'))$ is true.

principle of induction (standard form)

Conversely, suppose you are given a set X of natural numbers. Let $P(n)$ be the predicate $n \in X$. Then The Standard Principle of Induction applied to $P(n)$ can be phrased entirely in terms of the set X and the membership relation $\in$. When this is done, the result is precisely the Principle of Induction for Sets. So, the two principles are indeed just different formulations of the same thing and can be used interchangeably as needed.

A *proof by induction* involves a predicate such as $P(n)$, which one must prove to be true for every n. The Principle of Induction shows that this may be accomplished in two steps. The first, called the *base case*, involves verifying that $P(0)$ is true. This is often fairly easy. The second step, called the *inductive step*, involves proving an

implication $P(k) \implies P(k')$ for each k. As described in *The Predicate Calculus*, this amounts to choosing an arbitrary k, assuming that $P(k)$ is true, and then using this assumption to deduce that $P(k')$ is true. Please note: it is very important that in this step, the implication $P(k) \implies P(k')$ is proved *for an arbitrary, or general, natural number k, not simply for some particular natural number.*

We now present a few examples and exercises illustrating this method of proof. Many more examples and applications will occur in later parts of the text. In these examples and exercises we shall use the notation and operations that we are familiar with when working with numbers and algebra. In particular, we shall use $+$ in its usual sense and we shall use $n+1$ instead of n'.

Example: In this example, as in some others, we make use of standard arithmetic and algebraic facts about numbers for purposes of illustration. For n ranging over $\mathbb{N}$, let $P(n)$ be the predicate

$$0 + 1 + 2 + \ldots + n = \frac{n(n+1)}{2}.$$

We proceed by induction on n.

Base Case: When $n = 0$, both sides of the equation reduce to 0, so $P(0)$ is true.

Inductive Step: Let k be an arbitrary natural number, and assume that $P(k)$ is true. That is, $0 + 1 + 2 + \ldots + k = k(k+1)/2$. Add $k+1$ to both sides, obtaining a valid equality. Specifically, we get $0 + 1 + 2 + \ldots + (k+1)$ on the left side and $(k(k+1)/2) + (k+1) = (k+1)(k/2+1) = (k+1)(k+2)/2$ on the right. The equality of these two is exactly what $P(k+1)$ asserts, so $P(k+1)$ is true.

Therefore, by the Principle of Induction, $P(n)$ is true for all natural numbers n.

The above example is one of the first applications of induction that is presented in, say, a course on pre-calculus or calculus, so it should be well known to students.

There are many applications of induction of this sort. However, there are also many other types of applications, as you will see in the next example and throughout out development of the properties of $\mathbb{N}$. Here are a couple of examples of similar induction problems.

Exercise 4. Use standard facts about natural numbers that you have learned in earlier courses to prove the following by induction on n:

(a) For all n, $1 + 3 + 5 + \ldots + (2n + 1) = (n + 1)^2$.

(b) For all n, $0^3 + 1^3 + 2^3 + \ldots + n^3 = \big(n(n + 1)/2\big)^2$.

Example: Here, we give a completely different sort of example to illustrate the wide variety of problems that can be tackled by induction. This one is from geometry. Consider a plane in which some finite number of infinite straight lines have been drawn. They break up the plane into finitely many regions. We say that two such regions are *adjacent* if their boundaries share a positive-length segment of one of the lines. Now we are given a palette of m colors which we can use to color the regions, one color per region. If at most m colors are used in this way, so that no two adjacent regions have the same color, then we say that the line arrangement has an m-coloring. Prove that every line arrangement has a 2-coloring.

First, we reformulate this so that it can be viewed as an induction problem. Specifically, we suppose that there are n lines in the line arrangement. The predicate $P(n)$, then, is: Any arrangement of n lines has a 2-coloring.

Exercise 5. (a) Prove the base case, $P(0)$. (b) Choose an arbitrary k and assume $P(k)$. That is assume that any arrangement of k lines has a 2-coloring. Now consider any arrangement of $k+1$ lines. Remove one line, obtaining an arrangement of k lines.

Use $P(k)$ to obtain a 2-coloring of this arrangement, and then throw the removed line back in.

Prove that the 2-coloring you have can be modified to get a 2-coloring for the arrangement of $k+1$ lines. This will complete the induction step and thus conclude the proof by induction.

4. Definition by induction

definition by induction

As important and useful as proof by induction is, it is no more so than the method of definition by induction. Many students have seen instances of this in other courses: for example, raising a number to an arbitrary natural-number power, or forming factorials. Most such presentations are fine in the context of finite mathematics, but on the face of it they do not produce functions with domain the entire set of natural numbers $\mathbb{N}$. For example, $n!$ is often defined by the two equations $0! = 1$ and $(n+1)! = (n+1)n!$, which give a so-called recursive definition. This certainly defines $n!$ for as large n as we like. However, absent some further principle or argument, it only produces a finite set of function values. What we would like is for the factorial function to be a function $f : \mathbb{N} \to \mathbb{N}$, which is to say f should be a certain infinite set of ordered pairs in $\mathbb{N} \times \mathbb{N}$. We need a principle for producing such an f from the two equations given above.

Notice that we are in a position similar to what we encountered when constructing the sequence of natural numbers. There, we had an initial element and a rule of formation, but with these alone, we could not produce the totality of all natural numbers as a set in our universe $\mathcal{U}$. To obtain this totality, we required an additional assumption or axiom.

Now we could follow the same path here, by invoking a new principle that produces a general method of definition by induction. However, we want to keep our assumptions to a minimum. Moreover, it turns out that it is possible to derive a principle

sufficient for our purposes from what we have already assumed, i.e., from the Natural Number Axiom. We shall present this as a theorem. The proof of the theorem is fairly lengthy and would take us too far afield, so we omit that from these notes.

Theorem 1. *Suppose that we are given a set Y and a distinguished element $y_0 \in Y$.*
(a) Assume further that we are given a function $h : Y \to Y$.

definition of functions by induction

Then, there exists a unique function $f : \mathbb{N} \to Y$ satisfying: (i) $f(0) = y_0$, and (ii) $f(n') = h(f(n))$, for all natural numbers n.
(b) Instead of (a), assume we are given a sequence of functions $h_n : Y \to Y$, one for each natural number n.

Then, there exists a unique function $g : \mathbb{N} \to Y$ satisfying: (i) $g(0) = y_0$, and (ii) $g(n') = h_n(g(n))$, for all natural numbers n. $\square$

Since this theorem is fairly complex, we illustrate each of the two parts (a) and (b) with a specific definition by induction.

Example: Let a be a fixed real number, and let $h : \mathbb{R} \to \mathbb{R}$ be given by the equation $h(x) = ax$, for all real x. Let $1 \in \mathbb{R}$ be selected as the distinguished element. Then, by part (a) of the equation, there exists a unique function $f : \mathbb{N} \to \mathbb{R}$ such that $f(0) = 1$, and $f(n') = a \cdot f(n)$, for every natural number n. Clearly, $f(n) = a^n$, for all n. So, in this case, the theorem shows that we can define the entire"a to a power" function by induction.

Example: For each natural number n, define a function $h_n : \mathbb{N} \to \mathbb{N}$ by the rule $h_n(m) = n' \cdot m$, for every natural number m. Let $Y = \mathbb{N}$, and again choose 1 to be the distinguished element. Then, part (b) of the theorem produces a function $g : \mathbb{N} \to \mathbb{N}$ satisfying $g(0) = 1$, and $g(n') = h_n(g(n)) = n' \cdot g(n)$. Clearly, $g(n) = n!$,

for every natural number n. In this case, then, the theorem shows that we can define the entire factorial function by induction.

We shall be using definition by induction in an important way in the next section.

Exercise 6. In this exercise and the next, we use the standard properties and notation involving integers with which we are familiar from earlier courses.

For each $n \in \mathbb{N}$, we define a function $h_n : \mathbb{N} \to \mathbb{N}$ by the rule $h_n(x) = (2n + 3)x$. We use Theorem 1 to obtain a function $f : \mathbb{N} \to \mathbb{N}$ such that $f(0) = 1$ and $f(n') = h_n(f(n)$, for all n. Obtain a formula for $f(n)$. (The formula may include ellipses "..."). Do the same when h_n is given by $h_n(x) = (2n + 4)x$ and $f(0) = 2$.

Exercise 7. Consider the sequence of numbers

$$1, 2, 2^2, 2^{2^2}, 2^{2^{2^2}}, \ldots, \text{etc.}$$

with each number obtained by raising 2 to the preceding number. Use the above results to *prove* that one can define by induction a function $f : \mathbb{N} \to \mathbb{N}$, such that $f(n)$ is the n^{th} term in the above sequence (where we assume that the sequence starts with the 0^{th} term).

5. Addition

We have started with the most basic features of set theory and logic and, using only one additional axiom, the Natural Number Axiom, constructed a set, $\mathbb{N}$, that closely resembles the usual natural numbers. There is a so-called successor function σ defined on $\mathbb{N}$, which corresponds to our notion of counting, and we have seen that the simple concept of set inclusion leads to a well-ordering of $\mathbb{N}$ and the Principle of Induction. All of this is well and good, but there is a glaring gap. We have not

described how to do arithmetic with the natural numbers. And, while the counting features that we just described are probably the most fundamental, natural numbers without arithmetic would have comparatively little value. We now remedy this by *defining* the operation of addition; multiplication will come later.

Choose any $m \in \mathbb{N}$. We shall use induction to define a function $a_m : \mathbb{N} \to \mathbb{N}$, and we shall denote the function value $a_m(n)$ by $m + n$. To define a_m we make use of Theorem 1(a), and this requires us to have a function $h : \mathbb{N} \to \mathbb{N}$. In this case, the specification of h is easy: h is simply the successor function σ. We must also have a distinguished member of $\mathbb{N}$, which in this case, we specify to be m. Then Theorem 1(a) tells us that there exists a unique function $\mathbb{N} \to \mathbb{N}$, which we shall denote by a_m, such that

(a) $a_m(0) = m$.

(b) $a_m(n') = (a_m(n))'$.

Definition 1. For any m and n in $\mathbb{N}$, we denote $a_m(n)$ by $m + n$ and call it the *sum* of m and n (or the result of adding m to n, or m plus n).

In terms of the + notation, the two defining features of a_m may be written as follows:

(a) $m + 0 = m$.

(b) $m + n' = (m + n)'$.

It is easy to see directly from this definition that $0+0 = 0$, $1+1 = 1+0' = (1+0)' = 1' = 2$, $2+1 = 2' = 3$, $2+2 = 2+1' = (2+1)' = 3' = 4$, and so on, reproducing many of the elementary sums that we first learned in beginning arithmetic. However, to accomplish the important task of justifying the definition, we must show that it gives us the operation with which we are all familiar, i.e., that it has all the usual, general properties. Some of these are, indeed, immediate consequences of the definition.

The others can all be proved by induction. We sketch a few proofs by induction for illustration.

Lemma 4. Choose any natural numbers ℓ, m, and n. Then: (i) $m+0 = m = 0+m$; (ii) $m+1 = m' = 1+m$; (iii) $\ell + (m+n) = (\ell+m)+n$.

Proof. (i) $m + 0 = a_m(0) = m$, by definition, proving the first equality. The second equality, namely, $m = a_0(m)$ must be proved by induction on m. By definition, $a_0(0) = 0$, so the assertion is true for $m = 0$. Now assume it is true for $m = k$: that is, $k = a_0(k)$. Evaluate $a_0(k') = (a_0(k))' = k'$, the first equality following from property (b) of a_0 and the second from the induction hypothesis. This completes the proof of (i).

(ii) Evaluate $m + 1 = a_m(1) = a_m(0') = (a_m(0))' = m'$. (The reader should fill in the reasons for each equality.) The second equality, namely, $m' = a_1(m)$, is proved by induction. When $m = 0$, both sides reduce to 1, so the assertion is true for $m = 0$. Assume it for $m = k$, i.e., $k' = a_1(k)$, and then evaluate $a_1(k') = (a_1(k))' = (k')'$. Switching the order of the equality, we obtain $(k')' = a_1(k')$. But this is precisely the second equality in assertion (ii) of the lemma, with $m = k'$, as desired. Therefore, the induction proof of (ii), is complete.

(iii) Choose arbitrary ℓ and m, and let $P(n)$ be the predicate

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

We prove $(\forall n)P(n)$ by induction on n. $P(0)$ is the statement $\ell+(m+0) = (\ell+m)+0$. Both sides reduce to $\ell + m$, by what has already been proved, so $P(0)$ is true. Now

assume $P(k)$ for some k. We evaluate the left-hand side of $P(k')$:

$$\begin{aligned} \ell + (m + k') &= \ell + (m + k)' \\ &= \big(\ell + (m + k)\big)' \\ &= \big((\ell + m) + k\big)' \\ &= (\ell + m) + k'. \end{aligned}$$

Again, we leave it to the reader to supply reasons for each step. The resulting equality is precisely $P(k')$, so the inductive step is true. Therefore, the equality $\ell + (m + n) = (\ell + m) + n$ holds for every natural number n. Since ℓ and m were chosen arbitrarily in $\mathbb{N}$, the equality holds for all ℓ, m and n in $\mathbb{N}$. □

The reader may have noticed that statement (i) of the lemma is the familiar *additive identity rule* for arithmetic and that the third statement of the lemma is the familiar *associative law for addition*. Usually these are stated as axioms or properties that we assume. The lemma shows, however, that *these can be proved* from more rudimentary concepts.

The associative law is needed because addition is defined to be a *binary* operation on natural numbers. Once more than two numbers are to be added, we must decide how to group them so that we can do the addition two numbers at a time. In the case of three numbers, there are two ways of grouping the numbers (without changing their order)—or, as we may say, there are two ways of *associating* the numbers (into pairs). The associative law tells us that these two ways give equal sums.

Exercise 8. Show that there are five ways of associating four numbers into pairs, and then prove, using the associative law, that all five ways produce the same sum. (Do not use induction. Rather, use the already proved associative law as often as is needed.) How many ways of associating five numbers are there? List them.

This exercise shows that the associative law can be extended to the case of four numbers, and similar arguments extend the law to five numbers, etc. Indeed, there is a general inductive argument (which we do not go into in these notes) that extends the result to any number of numbers. This means that, for addition, we do not need to use parentheses to express arbitrarily long sums (e.g., $a_0 + a_1 + \ldots + a_n$), since any association into pairs yields the same result. So, from now on we feel free to omit parentheses in these situations.

Incidentally, earlier, we expressed some discomfort in using the notation "$n+1$" to represent the successor of n, since we had not yet defined addition. However, now, in light of assertion (ii) of the lemma, we are able to use either $n + 1$ or $1 + n$ in place of n'.

Exercise 9. Use the definition of addition and the method of proof by induction to prove the following statements:

(a) $(\forall m)(\forall n)(m + n = n + m)$. *commutative law for addition*

(b) $(\forall k)(\forall m)(\forall n)\big((m + k = n + k) \Rightarrow (m = n)\big)$. *cancellation law for addition*

The statements in the earlier lemma and in the exercise give the most important features of the additive algebra of the natural numbers, so we summarize them in the following theorem. Further algebra will come into play when we discuss multiplication later.

basic properties of addition

Theorem 2. *Let x, y, z be any natural numbers. Then the following hold:*

(a) $x + (y + z) = (x + y) + z$. *associativity of addition*

(b) $(x + z = y + z) \Rightarrow x = y$. *cancellation for addition*

(c) $x + 0 = 0 + x = x$. *identity law for addition*

(d) $x + y = y + x$. *commutativity of addition*

6. Addition and the natural ordering

There are important relationships between the natural ordering on $\mathbb{N}$ and the addition operation we have just defined. We have already seen an example of this, as expressed in terms of the successor function: namely, $m < n \implies m' < n'$. Using the $+$ notation, this becomes $m < n \implies m + 1 < n + 1$.

Exercise 10. Choose any m and n in $\mathbb{N}$ with $m < n$. Prove by induction on p that, for all $p \in \mathbb{N}, m + p < n + p$.

Exercise 11. Suppose that m, n, p and q are natural numbers such that $m+n = p+q$. Use the preceding exercise to prove that if $m < p$, then $n > q$.

Perhaps the most basic connection between the natural ordering and addition is given by the following fact:

Proposition 3. *Let m and n be any natural numbers. Then: $m \leq n \iff$ there exists a natural number k such that $n = m + k$.*

Proof. Assume that $m \leq n$. We shall show that $n = m + k$ for some k. Consider the set X consisting of all natural numbers that are either less than m, or of the form $m + k$, for some natural number k. Since $0 \leq m$, we have $0 \in X$, either because $0 < m$ or because $0 = m = m + 0$. Now suppose that some arbitrary natural number $\ell \in X$. If $\ell < m$, then $\ell' \leq m$, so we may argue for ℓ' just as we argued for 0 in the base case to conclude that $\ell' \in X$. If ℓ has the form $m + k$, for some k, then $\ell' = (m + k)' = m + k'$, which also has that form. So, in any case, $\ell' \in X$. It follows by the Principle of Induction for Sets that $X = \mathbb{N}$. Therefore, $n \in X$. Since, by assumption, n is not less than m, it must be the case that n has the form $m + k$ as required.

Now we prove the converse. Assume that $n = m + k$, for some k. If $n < m$, then by what we have just shown, $m = n + j$, for some j. Indeed, we can further conclude that this j is not equal to 0, since otherwise $m = n$ contrary to what we assume. Now combine the two equations: $n = m + k = (n + j) + k$, which can be written as $n + 0 = n + (j + k) = n + (k + j)$. By the Cancellation Law, it follows that $0 = k + j$. But, as we have seen, j is not 0, which means that j is a successor of some natural number, say $j = i'$. Therefore, $0 = k + i' = (k + i)'$, by the inductive definition of addition. This shows that 0 is a successor, which is clearly impossible. Therefore, $n < m$ is not true. Since $\leq$ is a linear ordering, it follows that $n \geq m$. □

For later use, it is convenient to rephrase this proposition in the following way: Let m and n be any natural numbers. The equation

$$m + x = n \quad \text{has a solution in } \mathbb{N} \text{ if and only if } \quad m \leq n.$$

Exercise 12. Prove the following addendum to the foregoing: If the equation has a solution in $\mathbb{N}$, then the solution is unique.

7. Variants on induction

Using the linear ordering of $\mathbb{N}$, it is now easy to formulate two variants on the usual method of induction.

shifted induction

Variant 1: This method does not have a commonly used name since it amounts to little more than a re-labeling. We might call it *shifted induction*, however, when we refer to it. This kind of induction arises when the predicate one would like to verify is not usefully defined for some small values of n. Perhaps $P(n)$ is defined only for

$n \geq$ some fixed number n_0 which depends on the particular problem. We can still do induction in this case, except that our initial value is no longer 0 but is n_0. The situation we are facing then is that $P(n_0)$ is known to be true and $(\forall k)\big(P(k) \Rightarrow P(k+1)\big)$ is known to be true, *provided* $k \geq n_0$. We would then like to be able to conclude that $P(n)$ is true for all $n \geq n_0$. Essentially, our entire frame of reference is shifted over by n_0 units.

This variant of induction is easy to derive from the standard induction principle via the following re-labeling trick. Define a new predicate Q by the formula $Q(n) = P(n_0 + n)$. Since, by Proposition 3, $k \geq n_0$ if and only if there is a natural number m such that $k = n_0 + m$, k ranges over all natural numbers $\geq n_0$ precisely when m ranges over all natural numbers. Therefore, the situation we confront can now be reformulated as follows: $Q(0)$ is true, and $(\forall m)\big(Q(m) \Rightarrow Q(m+1)\big)$ is true. So, ordinary induction implies that $Q(n)$ is true for all natural numbers n, and, hence, $P(n)$ is true for all natural numbers $n \geq n_0$.

Just as for ordinary induction, there is a version of shifted induction for sets. Thus, suppose that S is a set of natural numbers containing the natural number m and satisfying $k \in S \Rightarrow k+1 \in S$, for all natural numbers k. Then, *the principle of shifted induction for sets* concludes that $\{n : (n \in \mathbb{N}) \wedge (n \geq m)\} \subseteq S$. (Note that the hypotheses do not guarantee equality for general m. For example, $\mathbb{N}$ itself satisfies the hypotheses no matter what m is considered. In applications, one only needs the indicated inclusion, so this is not a problem.)

Variant 2: This method is sometimes called *strong induction* or *complete induction* even though it is equivalent to ordinary induction. In this form of induction, we still have the same initial step. But the induction step is replaced by the following

strong induction

$$(\forall k)\Big((\forall i)\big((i \leq k) \Rightarrow P(i)\big) \Rightarrow P(k+1)\Big).$$

That is, for all k, if $P(0), P(1), P(2), \ldots, P(k)$ are all true, then $P(k+1)$ is true. The initial step plus the induction step then imply the usual conclusion: $(\forall n)P(n)$ is true.

Exercise 13. Given a predicate $P = P(n)$, formulate a predicate Q such that by applying ordinary induction to Q, you get strong induction for P.

Of course, Variant 1 can be combined with strong induction: i.e., you don't have to start at 0 in strong induction.

Here is an example showing how strong induction, although logically equivalent to ordinary induction, is a convenient tool in proofs. We shall be referring to standard multiplicative properties of natural numbers, which have not yet been established in our logical development, but we are doing so only for illustration.

Proposition 4. *Every natural number $n > 1$ can be written as a product of primes.*

Proof. This proof is just an illustrative sketch. Since there are no natural numbers n satisfying $1 < n < 2$, we may start the strong induction at $n = 2$. When $n = 2$, the proposition is immediate, since 2 itself is a prime, so it is a one-fold product of primes. Suppose the proposition is true for all natural numbers $i \leq k$, and consider the natural number $k + 1$. If it is prime, we are done. If not, then by definition, it can be written as a product of two natural numbers $k + 1 = ab$, with neither a nor b equal to $k + 1$ or 1. Therefore, both a and b are $< k + 1$. Since there are no natural numbers strictly between k and $k + 1$, both a and b are $\leq k$, hence the strong induction hypothesis may be applied to them. They are thus both products of primes, and so, clearly their product is a product of primes. □

example of proof by strong induction

The point here is that since we are dealing with a multiplicative property, simply knowing the proposition for k may not be adequate for proving it for $k + 1$, since these two numbers are not related multiplicatively in a nice way. However, knowing

the proposition for all natural numbers $\leq k$ gives us the flexibility we need for the proof.

Strong induction also admits a version tailored for sets. It goes like this. Suppose that X is a set of natural numbers containing 0 such that, for each k,

$$(0, 1, \ldots k \in X) \implies k + 1 \in X.$$

Then $X = \mathbb{N}$. (There is also a version that combines this with shifted induction for sets, but we won't write this out explicitly.) The following exercise gives an important application of this:

Exercise 14. Prove that strong induction implies the well-ordering principle for the natural numbers. (Hint: Use the strong induction to prove the contrapositive of the well-ordering principle: that is, start with a set $Y \subseteq \mathbb{N}$ that has no smallest element, and prove that it is empty. Do this by proving that the complement of Y satisfies the conditions of strong induction for sets.)

This exercise, together with Proposition 2, shows that the Principle of Induction and the Well-Ordering Principle for the Natural Numbers are logically equivalent.

8. Multiplication

We proceed similarly to the case of addition, using the method of induction to define multiplication.

Choose any natural number m. We shall define a function $b_m : \mathbb{N} \to \mathbb{N}$ which is supposed to represent multiplication by m. So, the number m is held fixed throughout this definition process. To proceed, we again use Theorem 1(a), which requires the existence of a function $h : \mathbb{N} \to \mathbb{N}$. In this case, h is given by the formula $h(x) = x+m$. Note that we use addition in this definition, but that's okay because addition has

already been defined. Theorem 1 also requires us to choose a distinguished element in $\mathbb{N}$: our choice is the natural number 0.

Then, Theorem 1(a) tells us that there exists a unique function $b_m : \mathbb{N} \to \mathbb{N}$ satisfying

$$b_m(0) = 0, \text{ and}$$

$$b_m(n+1) = b_m(n) + m.$$

The first few values of b_m are:

$$b_m(0) = 0$$

$$b_m(1) = m$$

$$b_m(2) = m + m$$

$$b_m(3) = (m + m) + m$$

and so on. Clearly, b_m represents repeated addition of m, which coincides with our intuitive understanding of multiplication.

We write $b_m(n)$ as mn and call it the product of m by n or m times n. (Occasionally, we interpose a dot, as in $a \cdot b$, if this makes the product more readable.)

In terms of the product notation that we just introduced, the key defining properties of b_m can be written as

$$m \cdot 0 = 0, \text{ and}$$

$$m \cdot (n+1) = mn + m.$$

Exercise 15. Use induction on n to prove that for every natural number m, $0 \cdot m = 0$.

Exercise 16. Prove: If a and b are non-zero natural numbers, then $ab \neq 0$. (Hint: Use the fact that a and b are successors, together with the definition of multiplication, to prove that ab is a successor.)

Theorem 3. *For any natural numbers ℓ, m, and n, the following are true:*

basic properties of multiplication

(a) $\ell(m+n) = \ell m + \ell n$ *and* $(\ell + m)n = \ell n + mn$. *distributive laws*

(b) $(\ell m)n = \ell(mn)$. *associativity of multiplication*

(c) $\ell \cdot 1 = \ell = 1 \cdot \ell$. *multiplicative identity law*

(d) $\ell m = m\ell$. *commutativity of multiplication*

(e) *Suppose that* $\ell \neq 0$. *Then* $(\ell \cdot m \leq \ell \cdot n) \iff m \leq n$.

(f) *Suppose that* $\ell \neq 0$. *Then* $(\ell \cdot m = \ell \cdot n) \iff m = n$. *cancellation law for multiplication*

The listed properties are proved in the order listed, with each proof making use of results established earlier. We'll leave some of the proofs to the reader as exercises with the *caveat* that only the results proved earlier may be used.

Proof. : (a) Note that we already know one very special case of the first distributive law. Namely, we know that $\ell(m+1) = \ell \cdot m + \ell$, for every ℓ and m. For this is just the inductive part of the definition of multiplication. We use this fact a number of times in the induction proof that follows. First, we choose arbitrary ℓ and m, as in the proof of associativity for addition. We prove each distributive law for these ℓ and m by induction on n. The base case for the first distributive law asserts that $\ell(m+0) = \ell m + \ell \cdot 0$. The reader should check that both sides reduce to $\ell \cdot m$, and so they are equal. For the inductive step, we assume $\ell(m+k) = \ell \cdot m + \ell \cdot k$ and compute: $\ell(m+(k+1)) = \ell \cdot ((m+k)+1) = \ell(m+k) + \ell = (\ell \cdot m + \ell \cdot k) + \ell = \ell \cdot m + (\ell \cdot k + \ell) = \ell \cdot m + \ell \cdot (k+1)$, which completes the inductive step for the first distributive law, and concludes the induction proof for the chosen ℓ and m and any n. But, since ℓ and m were chosen arbitrarily, the result holds for all ℓ, m and n. The second law is proved

similarly. First, the base case: $(\ell+m)0 = 0 = 0+0 = \ell\cdot 0+m\cdot 0$. Next, the inductive step: $(\ell+m)(k+1) = (\ell+m)k+(\ell+m) = \ell\cdot k+mk+\ell+m = \ell\cdot(k+1)+m(k+1)$, which is what we needed to prove.

We have omitted giving reasons for the steps. It would be a good test of your understanding of the material if you can supply these.

Exercise 17. Prove item (b) of the theorem, the associative law for multiplication, by induction on n.

Exercise 18. Prove item (c) of the theorem, the multiplicative identity law. (Hint: for the first equality, do not use induction. Rather, use the fact that $1 = 0'$, together with the definition of multiplication. For the second equality, use induction on ℓ.)

We now continue the proof of the theorem.

(d) This is proved by induction on m. For the base case, we have $\ell \cdot 0 = 0 = 0 \cdot \ell$. For the inductive step, assume $\ell \cdot k = k \cdot \ell$, and compute $\ell \cdot (k+1) = \ell \cdot k + \ell = k\ell + \ell = k\ell + 1 \cdot \ell = (k+1)\ell$, completing the induction.

(e) Assume $m \leq n$. By Proposition 3, $n = m + k$, for some natural number k. Therefore, using item (a) above, the distributive law, we get $\ell n = \ell m + \ell k$. Applying Proposition 3 to this equality, we conclude that $\ell m \leq \ell n$.

We now prove the converse—namely $\ell m \leq \ell n \implies m \leq n$—by proving the contrapositive of this converse. That is, we show that $n < m \implies \ell n < \ell m$. (Make sure you understand why this is the contrapositive of the converse.) So, assume that $n < m$. Using Proposition 3 again, we get $m = n + j$, for some natural number j, which cannot be zero, since this would imply that $m = n$. Apply (a) to this equation, obtaining $\ell m = \ell n + \ell j$. It follows from Proposition 3 that $\ell n \leq \ell m$. But equality cannot hold in this case. For, if $\ell n = \ell m$, then from the previous equation and *additive* cancellation, it would follow that $0 = \ell j$, which contradicts the result of

Exercise 16, in light of the fact that both ℓ and j are non-zero. Therefore, we obtain the desired strict inequality $\ell n < \ell m$, completing the proof. $\square$

Exercise 19. Show how item (f) of Theorem 3 may be deduced from item (e).