4 The Natural Numbers

1. Numerals and Numbers

The next topic we consider is the set-theoretic reconstruction of the theory of natural numbers. This is a key part of the general program to reduce mathematics to set theory.

The basic strategy is to reduce classical arithmetic (thought of as the theory of the natural numbers) to set theory, and having done that proceed to reduce classical analysis (the theory of real numbers) to classical arithmetic. The remainder of classical mathematics (most importantly, Geometry) is then reduced to classical analysis.

In order to achieve the desired reduction, we must provide a set-theoretic definition of the natural numbers, as well as the standard arithmetic operations (addition, multiplication, etc.), and we must deduce the relevant theorems of arithmetic from the axioms of set theory.

This is a tall order! How do we start? How do we define the natural numbers? How do we define the number zero, the number one, etc.?

Besides numbers, there are numerals, which are the syntactic counterparts of numbers. Numerals are expressions like '0', '1', '2', etc., in the Arabic (decimal) numeral system, or expressions like '1001', '11011' in the binary numeral system, or expressions like 'I', 'II', 'III', etc., in the Roman numeral system. As we see below, we will use numerals in three different ways – as quantifiers, as predicates, and as singular-terms.

It is generally thought that the fundamental concept of number is quantitative, pertaining to the notion of "how many". As we see in the next chapter, this in turn is reducible to the relation "exactly as many", which in turn is reducible to "at least as many". In the current context, however, we regard the conceptually fundamental use of numbers to be as special quantifier concepts, akin to universal and existential quantifiers. Numbers as objects in their own right (the number 2, the number 3, etc.) are (logical or mental) constructions derived from the quantifier concepts.

We already have the universal and existential quantifiers. We can also add numerical quantifiers, as is familiar to students of elementary logic. For example, the sentence

there are *exactly two F*'s

can be rendered by the following formula:

$$\exists x \exists y(x \neq y \ \& \ \forall z(Fz \leftrightarrow .\ z=x \vee z=y))$$

We can give similar definitions for every possible number of F's. The following is an initial segment of the infinite series of official definitions of the numerical quantifiers.

(D1.0) $0\upsilon F[\upsilon] =_{df} \sim\exists\upsilon F[\upsilon]$
(D1.1) $1\upsilon F[\upsilon] =_{df} \exists x\forall\upsilon(F[\upsilon] \leftrightarrow \upsilon=x)$
(D1.2) $2\upsilon F[\upsilon] =_{df} \exists x\exists y(x\neq y \ \& \ \forall\upsilon(F[\upsilon] \leftrightarrow .\ \upsilon=x \vee \upsilon=y))$
(D1.3) $3\upsilon F[\upsilon] =_{df} \exists x\exists y\exists z(x\neq y \ \& \ y\neq z \ \& \ y\neq x \ \& \ \forall\upsilon(F[\upsilon] \leftrightarrow .\ \upsilon=x \vee \upsilon=y \vee \upsilon=z))$
etc.

Here, υ is any variable different from x,y,z, etc, and $F[\upsilon]$ is any formula in which variable υ occurs free.

With the *numerical-quantifiers* in hand, we can introduce a derivative series of *numerical-predicates*, which are intended to be applied to sets.

(D2.0) $0[A] =_{df} 0v[v \in A]$
(D2.1) $1[A] =_{df} 1v[v \in A]$
(D2.2) $2[A] =_{df} 2v[v \in A]$
etc.

Here, v does not occur free in A. For example, '$2[A]$' means that A has exactly two elements. These definitions may look circular, having numerals on both sides. However, the numerals on the right are quantifiers, and the numerals on the left are predicates. Furthermore, each numerical-quantifier can be expanded into primitive notation using (D1) above, which yields the following series of theorems.

(t1.0) $0[A] \leftrightarrow \sim\exists x[x \in A]$
(t1.1) $1[A] \leftrightarrow \exists x\forall y(y \in A \leftrightarrow y=x)$
(t1.2) $2[A] \leftrightarrow \exists x\exists y(x\neq y \,\&\, \forall z(z \in A \leftrightarrow .\, z=x \vee z=y))$
etc.

We also have the following theorems, which might be useful in seeing how the above numerical predicates work.

(T1.0) $\forall x[\, 0[x] \leftrightarrow x=\varnothing\,]$
(T1.1) $\forall x[\, 1[x] \leftrightarrow \exists y(x=\{y\})\,]$
(T1.2) $\forall x[\, 2[x] \leftrightarrow \exists y\exists z(y\neq z \,\&\, x=\{y,z\})\,]$
etc.

In other words, a set has zero elements iff it is identical to the empty set, a set has exactly one element iff it is identical to some singleton or other. And so on.

In our construction, we start with *numerals-as-quantifiers* – '$1x$', '$2x$', etc. Next, we construct *numerals-as-predicates* from numerals-as-quantifiers. The next step is to construct, by a further process of abstraction, *numerals-as-singular-terms*. This is the way numerals are used in arithmetic, as in sentences like '2+3=5'.

Under this construal of numerals as singular-terms, it is natural to think of numerals as names of objects – to wit, the numbers. How do we get from *numerals-as-predicates* to *numbers-as-objects*?

The process of abstraction is familiar. We have a predicate, e.g., '...is blue', and we construct a corresponding object "blueness" or "the property of being blue". This property has an extension, the class of all blue things. Indeed, in order to be conceptually economical, we might identify "blueness" with the class of all blue things.

To cite a more quantitative example, we begin with the predicate '... weighs 1 kilogram', and we identify the property of weighing 1 kilogram with the extension of this predicate, which is just the class of all things that weigh 1 kilogram.

Applying the above process of abstraction to the numerical predicates yields the following series of definitions of the numerical objects – the number zero, the number one, etc.

(d.0) $0 =_{df} \{x : 0[x]\}$
(d.1) $1 =_{df} \{x : 1[x]\}$
(d.2) $2 =_{df} \{x : 2[x]\}$
etc.

In other words, the number 0 is identified with the set of all empty sets, the number 1 is identified with the set of all singleton sets, etc.

The above definitions are very appealing, but unfortunately, they don't work. The problem is that, except for the first one, the predicates involved do not have proper extensions in our particular set theory. The extension of '0[x]' is the set of all empty sets, which is simply $\{\varnothing\}$. But the extension of '1[x]' is the set of all singletons, which is ill-defined, since its union is the universal set, a well-known non-set. Similarly with all the remaining numerical predicates.

So we go to Plan B. The property of being one meter long can be identified with the class of all one-kilogram objects. Alternatively, it can be identified with a particular one-kilogram object – the standard one-kilogram-object. This is in fact the way (weigh?) one-kilogram is defined – by reference to a standard one-kilogram object, housed in Paris. Until fairly recently, other measurement standard were ostensively defined, including one-meter. Advances in technology have made it possible to define one-meter by reference to a more reliable standard.

We can use the Paris kilogram as a model for defining the natural numbers. We cannot define 1 to be the set of *all* singletons, since this set is disallowed. So instead, we define 1 to be a particular "standard" singleton. Similarly, we define 2 to be a particular "standard" doubleton, etc.

So we have to make several choices; we have to:

(c0) choose a standard empty set
(c1) choose a standard singleton
(c2) choose a standard doubleton
etc.

The first choice is easy, since there is only one empty set – *the* empty set. This gives us our first definition.

(D2.0) $0 =_{df} \varnothing$

The remaining choices are more arbitrary, although in each case, there seems to be a natural choice.

For example, to choose the standard singleton, we choose some set A, and take its singleton $\{A\}$. What should A be? Looking above, we see the only set already in our construction is $\varnothing$, so it seems that the simplest choice is to set $A = \varnothing$. That yields our second definition.

(D2.1) $1 =_{df} \{\varnothing\}$

Next, we need a standard doubleton. This means finding two objects and forming their doubleton. We already have 0 and 1 (i.e., $\varnothing$ and $\{\varnothing\}$), so it seems natural to let these two sets constitute our standard doubleton. Once we have 0,1,2, we can collect them to form the number 3; once we have 0,1,2,3, we can form the number 4. And so on.

The choices we have made yield the following as our official definitions.

(D2.0) $0 =_{df} \varnothing$
(D2.1) $1 =_{df} \{0\}$
(D2.2) $2 =_{df} \{0,1\}$
(D2.3) $3 =_{df} \{0,1,2\}$
etc.

The underlying principle may be formulated as follows.

Every natural number is identical to the set of all of its predecessors.

0 has no predecessors (a set cannot have a negative number of elements!), so the set of all of 0's predecessors is the empty set. 1 has one predecessor, 0, so $1=\{0\}$. And so on.

Another way of looking at our construction is by way of the notion of *successor*, which is very important in later sections, and which is defined as follows.

(D3) $A^+ =_{df} A \cup \{A\}$

A^+ is called the *successor* of A. The definition applies to all sets, but its most useful application is to the natural numbers, as defined above. In particular, the successor of a number n is just the next number in the usual ordering. In other words, we have the following infinite series of theorems (exercise).

(T2.1) $1 = 0^+$
(T2.2) $2 = 1^+$
(T2.3) $3 = 2^+$
etc.

We conclude this section by noting that it is occasionally more useful to adopt the latter theorems as the definitions of the various numbers, in which case (D2.1), (D2.2), etc., become theorems.

2. The Axiom of Infinity

Having set-theoretically constructed all the natural numbers, the next obvious question is whether they can all be wrangled into a set – the set of all natural numbers. Curiously, the axioms we have so far postulated are inadequate to guarantee that such a set exists.

For this reason, we propose a further axiom, called the *Axiom of Infinity*, which goes as follows.

(AI) $\exists x(0 \in x \ \& \ \forall y(y \in x \rightarrow y^+ \in x))$

This axiom postulates the existence of a set that contains 0 and contains the successor of each of its elements. So it contains 0, 0^+, 0^{++}, 0^{+++}, and so on. The axiom of infinity is so called because it postulates the existence of a set with infinitely many elements (what 'infinitely many' means exactly is the topic of the next chapter).

In this connection, notice that we have postulated *infinitely many* sets, but up until now, we have not postulated any *infinitely large* sets. Now, we do that as well.

An alternative rendering of the Axiom of Infinity employs the notion of *successor set*, which is officially defined as follows.

(D4) $suc[A] =_{df} 0 \in A \ \& \ \forall x(x \in A \rightarrow x^+ \in A)$

In other words, A is a *successor set* iff A contains 0 and A contains the successor of any element of A.

This definition allows us to rewrite (AI) as follows.

(AI*) $\exists X(suc[X])$

In other words, there is at least one successor set.

The Axiom of Infinity says that *at least one* set contains all the natural numbers. It does not, however, say what else the set may contain in addition to the natural numbers. What we would like is a theorem like the following.

(t?) there is a set that contains 0,1,2, etc., **and nothing else**.

This set, if it exists, is none other than the set of natural numbers. How do we logically characterize the phrase 'and nothing else'?

The rigorous characterization of 'and nothing else' is important, not just to arithmetic, but to logic in general. For example, recall the way the syntax of propositional logic is defined. We say that every upper case letter is a formula, that the negation of any formula is a formula, that the conjunction of any two formulas is a formula, etc., and moreover nothing else is a formula. The usual way of characterizing the exclusion ('nothing else') clause is to say that the set of formulas is the *smallest set* of strings of symbols that contains all atomic formulas and is closed under the formation of negation, conjunction, etc.

Recall the definition of 'smallest' from an earlier chapter.

(d) $\sigma\upsilon F[\upsilon] =_{df} \iota\upsilon[F[\upsilon] \,\&\, \forall x(F[x] \rightarrow \upsilon \subseteq x)]$

Thus, to say that A is the smallest F is to say that A is an F, and A is included in every F.

Sometimes there is a smallest F; sometimes there isn't. Two examples might be useful. First, consider the following formula.

$$B \subseteq X \;\&\; C \subseteq X$$

In this example, to be an F is to include both B and C. So the smallest F is the smallest set that includes both B and C. Moreover, A is the smallest such set iff

$$B \subseteq A \,\&\, C \subseteq A \,\&\, \forall X(B \subseteq X \,\&\, C \subseteq X \;.\rightarrow\; A \subseteq X)$$

It is left to the reader to prove that this set does exist, that it is in fact just $B \cup C$; $B \cup C$ is the smallest set that includes both B and C.

As our second example, consider the following formula

$$X \neq \varnothing \;\&\; X \subseteq B$$

In this example, to be an F is be a non-empty subset of B. The question is whether there is a smallest such set. If B is itself empty, then it has no non-empty subsets, let alone a smallest one. If B is a singleton, $\{b\}$, then $\{b\}$ is the smallest non-empty subset of B. But if B has at least two elements a,b, then there is no smallest non-empty subset of B. The reason is that no singleton subset of B is included in any other singleton subset of B.

Sometimes there is a smallest F, sometimes there isn't. We next show that if there is *at least one* smallest F, then it is unique, which is a direct consequence of the Axiom of Extensionality. This can be alternatively described by saying that there is *at most one* smallest F.

Recall that 'at most one thing is a P' can be symbolized as follows.

$$\forall x \forall y(Px \,\&\, Py \,.\!\rightarrow x{=}y)$$

So, suppose both A and B are smallest F's. This means that:

$$F[A] \,\&\, \forall X(F[X] \rightarrow A \subseteq X) \;\;\&\;\; F[B] \,\&\, \forall X(F[X] \rightarrow B \subseteq X)$$

A little quantifier logic yields the following

$$A \subseteq B \;\;\&\;\; B \subseteq A$$

which, in virtue of Extensionality, yields the following.

$$A = B$$

Thus, the word 'the' in the phrase 'the smallest F' is justified, provided there is at least one smallest F.

Now, back to our original problem. What we want is the set that contains the natural numbers and nothing else. Claim: if there is such a set, then it is none other than *the smallest successor set*. The intuitive idea is that if a successor set contains any elements that are not natural numbers, then it is not the smallest successor set (exercise).

We dub the smallest successor set ω (omega), which is given by the following definition.

(D6) ω $=_{df}$ the smallest successor set.

Does such a set exist? The following is a sketch of the proof that it does. The details (including the lemmas) are left as an exercise.

(T) ω exists.

Proof :
By the Axiom of Infinity, there is at least one successor set; call it S. A successor subset of S is, by definition, a subset of S that is a successor set. Let C be the collection of all successor subsets of S; i.e., $C = \{X : X \subseteq S \;\&\; suc[X]\}$.

Claim: $\cap C$ is the smallest successor set.
Proving this divides into showing the following two:
(1) $\cap C$ is a successor set;
(2) $\cap C$ is included in every successor set.
Proof of (1): $\cap C$ is a successor set, since it is the intersection of a collection of successor sets (lemma 1).
In order to show (2), we first show the following:
Claim: $\cap C$ is the smallest element of C, which is to say it is the smallest successor subset of S.
Proof: we already showed that $\cap C$ is a successor set. It is also a subset of S, since it is the intersection of a collection of subsets of S (lemma 2). So $\cap C$ is itself a successor subset of S. Also, $\cap C$ is included in every element of C (lemma 3). Thus, $\cap C$ is a successor subset of S, and is included in every successor subset of S. So, by definition, $\cap C$ is the *smallest* successor subset of S.

Proof of (2): Suppose that B is a successor set, to show that $\cap C$ is included in B. Since suc[B] and suc[S], we have suc[$B \cap S$] (lemma 4). Now $B \cap S$ is a successor set, and $B \cap S$ is a subset of S, so it is a successor subset of S. But we have already shown that $\cap C$ is the smallest successor subset of S, so we must conclude that $\cap C \subseteq B \cap S$, but $B \cap S \subseteq B$, so $\cap C \subseteq B$, which was to be shown.

3. Mathematical Induction

We have gone to a lot of trouble to construct the set of natural numbers as the smallest successor set. In this section, we see why the trouble is worth it.

To say that ω is the smallest successor set is to say that

(t1) $suc[\omega] \ \& \ \forall X(suc[X] \rightarrow \omega \subseteq X]$

The second conjunct says that if X is a successor set, then every natural number is an element of X. This suggests a general and powerful strategy for proving theorems about the natural numbers of the form

(S) $\forall x(x \in \omega \rightarrow F[x])$

where $F[x]$ is any formula (with one free variable).

Suppose we want to show something of the form (S), which says (roughly, at least) that every natural number has property F. There are two cases to consider – F has a proper extension, or it doesn't. If F has a proper extension, which is to say that the set $\{x{:}F[x]\}$ is legitimate, then we need merely show that $\{x{:}F[x]\}$ is a successor set.

On the other hand, if F does not have a proper extension, we can form its "relative extension", as follows.

$$|F| = \{x : x \in \omega \ \& \ F[x]\},$$

whose existence is guaranteed by the Axiom of Separation. Now, our original formula (S) is logically equivalent to:

(S1) $\forall x(x \in \omega \rightarrow .\ x \in \omega \ \& \ F[x])$

In virtue of the Axiom of Separation and the definition of '$|F|$', showing (S1) amounts to showing

(S2) $\forall x(x \in \omega \rightarrow x \in |F|)$

which amounts to showing:

(S3) $\omega \subseteq |F|$

But, in light of the theorem (t1) above, in order to show (S3), it is sufficient to show

(S4) $suc[|F|]$

that is:

(S5) $0 \in |F| \ \& \ \forall x(x \in |F| \rightarrow x^+ \in |F|)$.

This naturally breaks into showing its two conjuncts,

1: $0 \in |F|$;
2: $\forall x(x \in |F| \rightarrow x^+ \in |F|)$

which, given the identity of $|F|$, amount to

1: $0 \in \omega \ \& \ F[0]$
2: $\forall x(x \in \omega \ \& \ F[x] \ .\rightarrow \ x^+ \in \omega \ \& \ F[x^+])$

But ω is a successor set, which means we have the following theorem.

(t2) $0 \in \omega \ \& \ \forall x(x \in \omega \rightarrow x^+ \in \omega)$.

So, given the resources of quantifier logic, we need actually only show the following.

1: $F[0]$
2: $\forall x(x \in \omega \ \& \ F[x] \ .\rightarrow \ F[x^+])$

Thus, in light of theorems (t1) and (t2), for any formula F (with or without a proper extension), we can show:

$\forall x(x \in \omega \rightarrow F[x])$

by showing

1: $F[0]$
2: $\forall x(x \in \omega \ \& \ F[x] \ .\rightarrow \ F[x^+])$

This procedure is called *Proof by Mathematical Induction*, and is one of the most powerful weapons in the logician's arsenal. This is because its use can be generalized from arithmetic to any domain of objects that are "inductively generated" (most importantly, well-formed formulas and proofs). For the moment, however, we are exclusively interested in the application of this proof technique to arithmetic.

The above proof-technique corresponds to the following theorem schema, which might be called the *Principle of Mathematical Induction*.

(PMI) $F[0] \ \& \ \forall x(x \in \omega \ \& \ F[x] \ .\rightarrow F[x^+]) \ .\rightarrow \ \forall x(x \in \omega \rightarrow F[x])$

The following theorem, which is similar to the second conjunct of (t1), is a logical consequence of (PMI) (although it is not an instance of (PMI)).

(t3) $\forall X(0 \in X \ \& \ \forall y(y \in \omega \ \& \ y \in X \ .\rightarrow \ y^+ \in X) \ .\rightarrow \ \omega \subseteq X)$

The outer proof of (t3) employs universal derivation (say, with constant 'A') and the inner proof employs (PMI), where the instantiating formula is '$x \in A$'. The details are left as an exercise.

4. Examples of Mathematical Induction

Before we turn to classical arithmetic, we consider a few simple applications of Mathematical Induction. The basic idea is that in order to show

(S) $\forall x(x \in \omega \rightarrow F[x])$

we show:

(BC) $F[0]$

and we show:

(IC) $\forall x(x \in \omega \ \& \ F[x] \ .\rightarrow \ F[x^+])$

The first part (BC) is called the *base case*, and the second part (IC) is called the *inductive case*. The latter is proven by universal conditional derivation, which means that we assume:

(IH) $m \in \omega \ \& \ F[m]$

to show:

(IS) $F[m^+]$

The former step is called the *inductive hypothesis*, and the latter step is called the *inductive step*.

Thus, Proof by Mathematical Induction can be schematized as follows.

	~~SHOW~~: $\forall x(x \in \omega \rightarrow F[x])$	
BC:	~~SHOW~~: $F[0]$	
IC:	~~SHOW~~: $\forall x(x \in \omega \ \& \ F[x] \ .\rightarrow \ F[x^+])$	UCD
IH:	$m \in \omega \ \& \ F[m]$	As
IS:	~~SHOW~~: $F[m^+]$	

Notice that it is customary to drop line (IC). Also notice that the first conjunct of (IH) is optional, in the sense that it might happen that one does not need to assume '$m \in \omega$' in order to show '$F[m^+]$'. It depends on the specific problem.

With the above proof schema in mind, let us do a few examples. We start by showing that every element of every natural number is itself a natural number, which is formally written as follows.

(T1) $\forall x(x \in \omega \rightarrow \forall y(y \in x \rightarrow y \in \omega))$

Given the way we have defined the natural numbers, as the series 0, {0}, {0,1}, {0,1,2}, etc., this is an intuitively obvious consequence, but its proof requires mathematical induction.

Before we continue, let us notice that (T1) is equivalent to:

(T1*) $\forall x(x \in \omega \rightarrow x \subseteq \omega)$.

Recall the definition of *transitive set*. A set is transitive iff every element is a subset. So, (T1) and (T1') are equivalent to the following theorem, which simply says that ω is a transitive set.

(T1**) trans[ω]

Both (T1) and (T1*) have forms that make proof by induction appropriate. We choose to prove (T1*) by mathematical induction, where the formula $F[x]$ is '$x \subseteq \omega$', as follows (Note: 'PT' means 'previous theorem').

	(1)	~~SHOW~~: $\forall x(x \in \omega \rightarrow x \subseteq \omega)$	MI
BC:	(2)	~~SHOW~~: $0 \subseteq \omega$	PT
IC:	(3)	~~SHOW~~: $\forall x(x \in \omega \,\&\, x \subseteq \omega \,.\!\rightarrow x^+ \in \omega)$	UCD
IH:	(4)	$m \in \omega \;\&\; m \subseteq \omega$	As
IS:	(5)	~~SHOW~~: $m^+ \subseteq \omega$	Def $\subseteq$
	(6)	~~SHOW~~: $\forall y(y \in m^+ \rightarrow y \in \omega)$	UCD
	(7)	$a \in m^+$	As
	(8)	~~SHOW~~: $a \in \omega$	SC
	(9)	$m^+ = m \cup \{m\}$	Def $^+$
	(10)	$a \in m \cup \{m\}$	7,9,IL
	(11)	$a \in m \vee a=m$	10, Def $\cup$
	(12)	c1: $a \in m$	As
	(13)	$m \subseteq \omega$	4b
	(14)	$a \in \omega$	12,13, PT
	(15)	c2: $a=m$	As
	(16)	$a \in \omega$	4a, IL

This shows that the set ω of all natural numbers is transitive. We can also show that every element of ω (i.e., every natural number) is transitive, which is the following theorem.

(T2) $\forall x(x \in \omega \rightarrow \text{trans}[x])$
$\forall x(x \in \omega \rightarrow \forall y(y \in x \rightarrow y \subseteq x))$

We also prove this by induction, where the formula $F[x]$ is '$\forall y(y \in x \rightarrow y \subseteq x)$'.

	(1)	~~SHOW~~: $\forall x(x \in \omega \rightarrow \forall y(y \in x \rightarrow y \subseteq x))$	MI
BC:	(2)	~~SHOW~~: $\forall y(y \in 0 \rightarrow y \subseteq 0)$	PT
IC:	(3)	~~SHOW~~: $\forall x(x \in \omega \,\&\, \forall y(y \in x \rightarrow y \subseteq x) \rightarrow \forall y(y \in x^+ \rightarrow y \subseteq x^+)$	UCD
IH:	(4)	$m \in \omega \;\&\; \forall y(y \in m \rightarrow y \subseteq m)$	As
IS:	(5)	~~SHOW~~: $\forall y(y \in m^+ \rightarrow y \subseteq m^+)$	UCD
	(6)	$a \in m^+$	As
	(7)	~~SHOW~~: $a \subseteq m^+$	SC
	(8)	$a \in m \cup \{m\}$	6, Def $^+$
	(9)	$a \in m \vee a=m$	8, Def $\cup$
	(10)	c1: $a \in m$	As
	(11)	$a \subseteq m$	4b, QL
	(12)	$m \subseteq m^+$	Lemma A
	(13)	$a \subseteq m^+$	11,12,PT
	(14)	c2: $a=m$	As
	(15)	$m \subseteq m^+$	Lemma A
	(16)	$a \subseteq m^+$	14,15,IL

[Lemma A is quite trivial: $m \subseteq m^+$, since $m^+ = m \cup \{m\}$, and $m \subseteq m \cup \{m\}$.]

We conclude this section by proving that a natural number m is not an element any element of m, which is to say

(T3) $\forall x(x \in \omega \rightarrow \forall y(y \in x \rightarrow x \notin y))$

This has a form that makes mathematical induction an appropriate proof technique. In this case, the induction formula $F[x]$ is '$\forall y(y \in x \rightarrow x \notin y)$'. The proof goes as follows.

	(1)	~~SHOW~~: $\forall x(x \in \omega \rightarrow \forall y(y \in x \rightarrow x \notin y))$	MI
BC:	(2)	~~SHOW~~: $\forall y(y \in 0 \rightarrow. \; 0 \notin y)$	PT
	(3)	trivial, since nothing is an element of 0, which is $\varnothing$	
IC:	(4)	~~SHOW~~: $\forall x(x \in \omega \;\&\; \forall y(y \in x \rightarrow x \notin y) \rightarrow \forall y(y \in x^+ \rightarrow. \; x^+ \notin y)$	UCD
IH:	(5)	$m \in \omega \;\&\; \forall y(y \in m \rightarrow m \notin y)$	As
IS:	(6)	~~SHOW~~: $\forall y(y \in m^+ \rightarrow. \; m^+ \notin y)$	UCD
	(7)	$a \in m^+$	As
	(8)	~~SHOW~~: $m^+ \notin a$	ID
	(9)	$m^+ \in a$	As
	(10)	~~SHOW~~: $\times$	SC
	(11)	$a \in m \vee a=m$	7, Def $^+$,$\cup$
	(12)	~~SHOW~~: $m^+ \in m$	SC
	(13)	c1: $a \in m$	As
	(14)	$a \subseteq m$	5a,13,T2
	(15)	$m^+ \in m$	3,7
	(16)	c2: $a=m$	As
	(17)	$m^+ \in m$	3,9
	(18)	$m^+ \subseteq m$	5a,15,T2
	(19)	$m \in m^+$	Lemma B
	(20)	$m \in m$	11,12
	(21)	$\times$	5b,20,QL

In the above proof, T2 is the previous theorem that every natural number is transitive. Lemma B is a simple consequence of the definition of $^+$ and $\cup$: $m \in \{m\} \subseteq m \cup \{m\} = m^+$.

The following is slightly stronger claim, which can also be proven by induction.

(T4) No natural number is a subset of any of its elements.

$$\forall x(x \in \omega \rightarrow \forall y(y \in x \rightarrow \sim[x \subseteq y])$$

Its proof is left as an exercise. Also, show that (T3) follows logically from (T2) and (T4).

5. The Peano Axioms

We have now defined all the numbers and we have postulated a set to contain them all. What remains to be done is to show how classical arithmetic can be deduced from set theory using these definitions. It is accordingly of considerable importance to establish exactly what classical arithmetic is. For this reason, we make a digression out of set theory proper, to talk about the classical theory of arithmetic.

Classical arithmetic was given its definitive formulation by Peano (and, independently, by Dedekind). Like classical geometry, classical arithmetic involves a few undefined (primitive) notions; in the case of arithmetic, the primitive notions include the following.

(u1) 0 (zero)
(u2) ...is a number
(u3) the successor of...

The postulates of classical arithmetic, which are called the Peano Postulates (Axioms), may be written as follows.

(p1) 0 is a number;
(p2) the successor of any number is a number;
(p3) 0 is not the successor of any number;
(p4) no two numbers have the same successor;
(p5) if 0 has a property P, and if the successor of a number x has P whenever x has P, then every number has P.

What remains is to formulate classical arithmetic in a formal language. For example, we can formulate it in a first order language with the following non-logical symbols.

(s1) $N[x]$ :: x is a number
(s2) $s(x)$ :: the successor of x
(s3) 0 :: the number zero

The Peano Axioms are then formulated as follows.

(a1) $N[0]$
(a2) $\forall x(N[x] \rightarrow N[s(x)]$
(a3) $\sim\exists x[N[x] \,\&\, s(x)=0]$
(a4) $\forall x \forall y(N[x] \,\&\, N[y] \,\&\, s(x)=s(y) \;.\rightarrow\; x=y)$
(a5) $F[0] \,\&\, \forall x(N[x] \,\&\, F[x] \;.\rightarrow\; F[s(x)]) \;.\rightarrow\; \forall x(N[x] \rightarrow F[x])$

Notice carefully that (a5) is an *axiom schema*; it is short for infinitely many axioms, one for each formula F. In this respect, (a5) is like the Axiom of Separation.

Actually, in *pure* arithmetic, the predicate 'N[]' is unnecessary, much as the predicate 'is a set' is unnecessary in *pure* set theory. In pure arithmetic, the domain of discourse is delimited from the outset, so that '$\forall$x' means 'for any number x', just as in pure set theory, '$\forall$x' means 'for any set x'.

So, if we are doing pure arithmetic, adopting the quantifier-domain convention, we can rewrite the Peano Axioms much more succinctly, as follows.

(a3) $\sim\exists x[s(x)=0]$
(a4) $\forall x\forall y(s(x)=s(y) \ .\rightarrow \ x=y)$
(a5) $F[0] \ \& \ \forall x(F[x] \rightarrow F[s(x)]) \ .\rightarrow \ \forall xF[x]$

Notice the total disappearance of (a1) and (a2).

This is the *first-order formulation* of arithmetic. There is also a *second-order formulation*. The difference between the two concerns only postulate (a5). Second order logic permits quantification over predicates in addition to quantification over singular-terms. It is accordingly considerably more powerful than first order logic. In place of the axiom *schema* (a5), we can write down a single axiom, as follows.

(a5*) $\forall P(P[0] \ \& \ \forall x(P[x] \rightarrow P[s(x)]) \ .\rightarrow \ \forall xP[x]$

Second order logic is powerful, but unfortunately it is *incomplete*; it cannot be axiomatized; it admits no adequate derivation system. In particular, for any system of deduction, the second-order Peano Postulates have logical consequences that cannot be deduced in that system!

By contrast, first order logic is complete; it can be axiomatized; it admits a derivation system. For example, if a first order argument is valid, its conclusion can be deduced from its premises inside *our* derivation system.

Remember, argument-validity is defined by reference to interpretations: an argument is valid iff no interpretation makes the premises true and the conclusion false. A simple example, in propositional logic, validity is defined by reference to truth-tables. In any case, validity is not defined by reference to derivations.

A remaining question is whether validity can be characterized by a derivation system. Our derivation system adequately characterizes first order logic, although this nifty fact has never been proven. The adequacy of derivation systems is a major issue in metalogic.

So far, we have not argued why classical arithmetic is founded on the Peano Axioms. This is primarily a matter of showing that all the familiar arithmetic "laws" can be deduced from them. This includes most prominently the laws of addition, multiplication, etc.

We postpone these matters, however, until after we reformulate classical arithmetic inside set theory.

6. The Set-Theoretic Formulation of the Peano Axioms

Before dealing with the details of arithmetic, we need to convince ourselves that the Peano Postulates can be deduced from the axioms of set theory, using the definitions proposed earlier in this chapter. This turns out to be fairly easy given what we have already done in Sections 3 and 4.

First of all, the following is the set-theoretic reconstruction of the three primitive notions of classical arithmetic. Notice that, since we are *not* doing *pure* arithmetic here, we do need the predicate '...is a number'.

(1) $N[x]$ :: $x \in \omega$
(2) $s(x)$ :: x^+ $(= x \cup \{x\})$
(3) 0 :: $\varnothing$

The Peano Postulates (first order form) are then reformulated as follows.

(a1) $0 \in \omega$
(a2) $\forall x(x \in \omega \rightarrow x^+ \in \omega)$
(a3) $\sim\exists x[x \in \omega \ \& \ x^+ = 0]$
(a4) $\forall x\forall y(x \in \omega \ \& \ y \in \omega \ \& \ x^+ = y^+ \ .\rightarrow \ x=y)$
(a5) $F[0] \ \& \ \forall x(x \in \omega \ \& \ F[x] \ .\rightarrow F[x^+]) \rightarrow \forall x(x \in \omega \rightarrow F[x])$

First, observe that (a1) and (a2) merely state that ω is a successor set, which follows immediately from the definition of ω. Next, (a3) is actually a consequence of the following stronger theorem.

(a3+) $\sim\exists x[x^+ = 0]$

> **Proof:** To prove (a3+), suppose $a^+=0$, which is to say $a^+=\varnothing$, where $a^+=a\cup\{a\}$. Now, $a\in\{a\}$, so $a\in a\cup\{a\}$, so $a\in\varnothing$, which contradicts an earlier theorem that $\varnothing$ has no elements.

Before looking at (a4), whose proof requires a little more work, we look at (a5), which we immediately observe to be identical to the Principle of Mathematical Induction, which we have already proved as a consequence of the definition of ω as the smallest successor set.

Finally, we examine (a4). To prove (a4), suppose $m,n\in\omega$ and $m^+=n^+$, to show $m=n$. Furthermore, suppose $m\neq n$, to show a contradiction. Now, $m\in m^+$, so $m\in n^+$, so $m\in n\vee m=n$. The latter is ruled out, by hypothesis, so we have $m\in n$. Parallel reasoning shows that $n\in m$. The conjunction '$m\in n\&n\in$ m' contradicts our earlier theorem that no number is an element of any of its elements.

Thus, we see that the Peano Postulates, as formulated inside set theory, can be deduced from the axioms of set theory. It remains to be shown that the usual principles of arithmetic, concerning addition, multiplication, etc., can also be deduced.

7. The Recursion Theorem; Definition by Induction

In order to do arithmetic as it is ordinarily construed, we must define the usual arithmetic operations (addition, multiplication, exponentiation, etc.), and we must deduce the laws of arithmetic (commutativity, associativity, etc.) from these definitions together with the postulates of set theory.

The definition of addition, multiplication, etc., is a special case of a general technique of definition, called *definition by induction*.

The basic idea of definition by induction is as follows. We begin with a set A, and we pick out an *initial* (or base, or root) *element* a, then we select a *generating function* g on A (i.e., g maps A into A). Given these three things, A,a,g, we define a function, f, by the following infinite series of clauses.

(0) $f(0) = a$
(1) $f(1) = g(a)$
(2) $f(2) = g(g(a))$
(3) $f(3) = g(g(g(a)))$
etc.

This can be more succinctly described as dividing into two clauses, the *base clause*, and the *inductive clause* (which are quite reminiscent of proof by induction).

(BC) $f(0) = a$
(IC) $f(x^+) = g(f(x))$

This says we pick out an initial element a in A and decree it to be $g(0)$. Next, we construct (generate) $f(1)$, $f(2)$, etc. by successive application of the generating function g. To get $f(1)$, we apply g to a; to get $f(2)$, we apply g to $f(1)$; to get $f(3)$, we apply g to $f(2)$; and so forth. All we need is an initial base object, a, and a generating function g, and we are off to the races.

The obvious question is whether such hocus-pocus works. It is certainly different from any definitional scheme we have seen before. Its coherence ultimately depends on our being able to rewrite *inductive definitions* as *explicit definitions* (of the form "$e_1 =_{df} e_2$".

The associated theorem is customarily called the *Recursion Theorem*, which is customarily stated as follows.

(RT) For any set A, for any element a of A, and for any function g on A, there exists a function f from ω to A such that:

(1) $f(0) = a$
(2) $\forall x[x \in \omega \rightarrow f(x^+) = g(f(x))]$

In proving the Recursion Theorem, one proves that the hypothesized function f is identical to the function *explicitly* defined as follows; here, fun(A,a,g) is the function inductively defined using set A, element a, and generating function g.

(D) $\text{fun}(A,a,g) =_{df} \cap\{R: R \subseteq \omega \times A \;\&\; 0Ra \;\&\; \forall x \forall y[xRy \rightarrow x^+Rg(y)]\}$

8. Two-Place Functions as Families of One-Place Functions

Definition by induction (which is based on the Recursion Theorem) can be used to define the familiar arithmetic operations – addition, multiplication, and exponentiation.

All three operations are two-place functions on the set ω of natural numbers. Definition by induction defines one-place functions, so we need first to examine how a two-place function can be defined as a family of one-place functions.

In order to clarify the intuition involved, we work backwards for a moment. Suppose we have a two-place function f on set A, which is to say that f assigns elements of A to two-tuples of elements of A (i.e., $f{:}A^2 \to A$). First, we adopt the usual abbreviation.

(d1) $f(a,b) =_{df} f(\langle a,b \rangle)$

Now, there is a corresponding family f^*, indexed by A, of (one-place) functions on A

$$\langle f^*_i{:}\ i \in A \rangle$$

defined so that each term f^*_i is a function defined so that

$$f^*_i(x) = f(i,x)$$

for every x in A.

Indeed, to a large extent, the two expressions are mere notational variants of one another, much as 'f_i' is a notational variant of '$f(i)$'. Of course, this is not strictly true. Whereas f is a two-place function on A, f^* is a function that assigns a one-place function on A to each element of A. This is precisely what it means to say that, f^* is a family, indexed by A, of functions on A.

The other direction also works: every A-family (family indexed by A) of functions on A corresponds to a two-place function on A. In particular, given an A-family $\langle f_i{:}\ i \in A \rangle$ of functions on A, we define the corresponding two-place function, f^*, in the obvious manner; the first expression is the official definition; the second expression is a convenient abbreviation.

(d2) $f^* =_{df} \iota f(f{:}A^2 \to A \ \&\ \forall xy(x,y \in A \to f(x,y) = f_x(y))$

(d3) $f^*(x,y) =_{df} f_x(y)$

In the case of arithmetic, the intuition is that we take a two-place function on ω, and we decompose it into an infinite series (an ω-sequence) of one-place functions. For example, the function adding-x-and-y is decomposed into the series of functions, adding-0-and-y, adding-1-and-y, adding-2-and-y, etc.

9. The Familiar Arithmetic Operations

Having seen how every two-place function can be identified with a family of one-place functions, we now proceed to the definition of the familiar arithmetic operations.

The Recursion Theorem says that for any set A, and element a in A, for any function g on A, there is a function f from ω into A such that $f(0) = a$, and $f(n^+) = g(f(n))$, for all n.

We are interested in functions from ω into ω, so we set $A = \omega$. We next need an initial element, and a generating function. The most obvious function to start with is the *successor function* on ω, which is defined as follows.

(d1) $suc =_{df} \{(x,x^+): x \in \omega\}$

The Recursion Theorem then tells us that for any initial element m we pick, there is a function, call it s_m, with the following properties.

(BC) $s_m(0) = m$

(IC) $s_m(n^+) = [s_m(n)]^+$

This defines an ω-sequence of one-place functions on ω. We define the corresponding two-place function on ω, also denoted s, in the usual manner.

(d2) $s(m,n) =_{df} s_m(n)$

What function is s? The following are some examples of its application.

$s(3,0) = s_3(0) = 3$

$s(3,1) = s_3(1) = s_3(0^+) = [s_3(0)]^+ = 3^+ = 4$

$s(3,2) = s_3(2) = s_3(1^+) = [s_3(1)]^+ = 4^+ = 5$

$s(6,0) = s_6(0) = 6$

$s(6,1) = s_6(1) = s_6(0^+) = [s_6(0)]^+ = 6^+ = 7$

$s(6,2) = s_6(2) = s_6(1^+) = [s_6(1)]^+ = 7^+ = 8$

The reader has probably surmised that $s(m,n)$ is just $m+n$. We have succeeded in defining addition!

The inductive definition of addition is in fact in complete agreement with the intuitive definition based on counting. To add n to m, you start at m and count n steps and then stop, which is to say that you apply the successor operation to m n times. The successor function is the set-theoretic explication of counting (the process of going from one number to the next one).

We have thus seen that addition can be defined as a family of functions $\langle s_m: m \in \omega\rangle$, where each term s_m is the function inductively generated using m as initial element and using the successor function as generating function g. Using earlier notation, used to convert inductive definitions into explicit definitions, we can summarize the above as follows.

(D7) $s_m =_{df} \text{fun}(\omega,m,suc)$

The next function on our agenda is multiplication. In elementary school, multiplication is defined in terms of addition; in order to multiply m times n, one adds m together n times. The underlying intuition is summarized by the following infinite list.

(m0) $m \times 0 = 0$
(m1) $m \times 1 = m$
(m2) $m \times 2 = m+m$
(m3) $m \times 3 = m+m+m$
(m4) $m \times 4 = m+m+m+m$
etc.

The formal definition differs from this informal definition only in that it is sufficiently rigorous to allow formal proofs pertaining to all natural numbers, not just the first few.

In giving the rigorous definition of multiplication, we use the Recursion Theorem. This time we set the initial object to be 0, and we use s_m as our generating function, so as to obtain a family $\langle p_m : m \in \omega \rangle$ of functions, one for each number m. Each p_m is defined inductively as follows.

(BC) $p_m(0) = 0$

(IC) $p_m(n^+) = s_m(p_m(n))$

The explicit definition is given as follows.

(D8) $p_m =_{df} \text{fun}(\omega, 0, s_m)$

Next, we define the corresponding two-place operation, and we introduce the customary notation, as follows.

(d3) $p(m,n) =_{df} p_m(n)$

(d4) $m \times n =_{df} p(m,n)$

If we use the more customary notation ('+' and '×') then we can write the inductive clause as follows.

(IC*) $m \times (n^+) = m + (m \times n)$

The following are examples of the application of the definition.

$3 \times 1 = p_3(1) = p_3(0^+) = s_3(p_3(0)) = 3+0 = 3$
$3 \times 2 = p_3(2) = p_3(1^+) = s_3(p_3(1)) = 3+(3+0) = 3+3$
$3 \times 3 = p_3(3) = p_3(2^+) = s_3(p_3(2)) = 3+(3+(3+0)) = 3+3+3$

In other words, the inductive definition of multiplication indeed defines multiplication in terms of addition in the familiar manner.

After multiplication comes exponentiation (raising a number to a power – squaring, cubing, etc.) The intuitive definition is provided by the following infinite series. We write '$m \uparrow n$' to express the number m raised to the nth power. The notation is not entirely standard, but it does make more sense after we see that ordinary exponentiation is just the first term in an infinite series of hyper-exponentiation functions.

(e0) $m\uparrow 0 = 1$
(e1) $m\uparrow 1 = m$
(e2) $m\uparrow 2 = m\times m$
(e3) $m\uparrow 3 = m\times m\times m$
(e4) $m\uparrow 4 = m\times m\times m\times m$
etc.

The rigorous definition appeals to the Recursion Theorem. In this case, we set 1 as the initial element, and we use p_m as our generating function, so as to obtain a family of functions $\langle e_m\colon m \in \omega\rangle$. The inductive definition is given as follows.

(BC) $e_m(0) = 1$

(IC) $e_m(n^+) = p_m(e_m(n))$

The explicit definition goes as follows.

(D9) $e_m =_{df} \text{fun}(\omega,1,p_m)$

Writing '$a\times b$' in place of '$p_a(b)$', (IC) can be rewritten as

(IC*) $e_m(n^+) = m \times e_m(n)$

We next introduce notation for the corresponding two-place function, as follows.

(d5) $m\uparrow n =_{df} e_m(n)$

We conclude this section by reviewing the various definitions.

(d) $0 =_{df} \varnothing$

(d) $a^+ =_{df} a \cup \{a\}$

(d) $suc[A] =_{df} 0 \in A \ \&\ \forall x(x \in A \rightarrow x^+ \in A)$

(d) $\omega =_{df} \sigma x\ \{\ suc[x]\ \}$

(d) $fun(A,a,g) =_{df} \iota f\{f\colon \omega\rightarrow A \ \&\ f(0)=a \ \&\ \forall x[f(x^+) = g(f(x))]\}$

(d1) $suc =_{df} \{(x,x^+)\colon x \in \omega\}$

(D7) $s_m =_{df} \text{fun}(\omega,m,suc)$

(D7') $m+n =_{df} s_m(n)$

$m+0 = m$
$m+n^+ = (m+n)^+$

(D8) $p_m =_{df} \text{fun}(\omega,0,s_m)$

(D8') $m\times n =_{df} p_m(n)$

$m \times 0 = 0$
$m \times n^+ = m + (m \times n)$

(D9) $e_m =_{df} \text{fun}(\omega, 1, p_m)$

(D9') $m \uparrow n =_{df} e_m(n)$

$m \uparrow 0 = 1$
$m \uparrow n^+ = m \times (m \uparrow n)$

10. Hyper-Exponentiation

We have now exhausted the *usual* arithmetic functions. There is, of course, subtraction and division, but these are more properly part of the theory of integers and rational numbers, respectively, rather than part of arithmetic. In particular, neither subtraction nor division is a function on ω (e.g., neither 2-3 nor 2/3 is a natural number). Accordingly, they are not proper arithmetic operations.

Nevertheless, there are infinitely many more proper arithmetic operations, not usually considered in elementary arithmetic.

Recall what has transpired so far. Arithmetic is founded on the notion of counting, which is formulated by the successor function. We define addition in terms of successor, multiplication in terms of addition and successor, and exponentiation in terms of multiplication and successor.

We have only scratched the surface! There is, in fact, an infinite series of arithmetic operations, of which the first three are only the beginning.

I introduce the term 'hyper-exponentiation' for the remaining functions. The idea is to form an infinite series of functions, starting with ordinary exponentiation. Each one is defined in terms of the previous one in the same way that exponentiation is defined in terms of multiplication and multiplication is defined in terms of addition.

(m) $m \times n = m+m+...+m$
(e1) $m \uparrow n = m \times m \times ... \times m$
(e2) $m \uparrow\uparrow n = m \uparrow m \uparrow ... \uparrow m$
(e3) $m \uparrow\uparrow\uparrow n = m \uparrow\uparrow m \uparrow\uparrow ... \uparrow\uparrow m$
etc. +----------+ ◄—— n times

So far, so good, except for one slight difficulty. What does $2 \uparrow 2 \uparrow 3$ equal? Addition and multiplication are associative, which means that parentheses are optional. Not so exponentiation, for

$$2 \uparrow (2 \uparrow 3) = 2 \uparrow 8 = 256,$$

whereas

$$(2 \uparrow 2) \uparrow 3 = 4 \uparrow 3 = 64.$$

The intuitive definitions above don't even work intuitively unless we say precisely what '$m \uparrow m \uparrow m$' means. The choice is clear given our later inductive definition.

$$\begin{array}{lll} m\uparrow m\uparrow m & =_{df} & m\uparrow(m\uparrow m) \\ m\uparrow m\uparrow m\uparrow m & =_{df} & m\uparrow(m\uparrow(m\uparrow m)) \\ m\uparrow m\uparrow m\uparrow m\uparrow m & =_{df} & m\uparrow(m\uparrow(m\uparrow(m\uparrow m))) \\ \text{etc.} & & \end{array}$$

So for example,

$$\begin{array}{l} 2\uparrow 2\uparrow 2 = 2\uparrow(2\uparrow 2) = 2\uparrow 4 = 16 \\ 2\uparrow 2\uparrow 2\uparrow 2 = 2\uparrow 16 = 65{,}536 \\ 2\uparrow 2\uparrow 2\uparrow 2\uparrow 2 = 2\uparrow 65{,}536 = ?? \end{array}$$

This is made more precise in the following inductive definitions.

For each number m, we define the function E^2_m as follows.

(BC) $E^2_m(0) = 1$

(IC) $E^2_m(n^+) = E_m(E^2_m(n))$

Here, $E_m(n) = m\uparrow n$, and $E^2_m(n) = m\uparrow\uparrow n$.

Similarly, we define the next hyper-exponentiation function, E^3_m, in terms of E^2_m, as follows.

(BC) $E^3_m(0) = 1$

(IC) $E^3_m(n^+) = E^2_m(E^3_m(n))$

By now the reader may have noticed that we have the makings of a "super" inductive definition of a sequence

$$\langle E^n_m : n \in \omega \rangle$$

of hyper-exponentiation functions, one such sequence for each number m. Term number 1 is ordinary exponentiation. Moreover, term number 0 is multiplication. The following is the complete inductive definition, for specific number m.

(BC) $E^0_m = P_m$

(IC) $E^n_{m+} = \text{fun}(\omega,1,E^n_m)$

Having made the formal inductive definition, we re-introduce the binary operator notation as follows.

(d1) $m\uparrow n =_{df} E^1_m(n)$

(d2) $m\uparrow\uparrow n =_{df} E^2_m(n)$

(d3) $m\uparrow\uparrow\uparrow n =_{df} E^3_m(n)$

etc.

We conclude this section by noting that hyper-exponentiation affords notation considerably more compact than scientific notation (ordinary exponentiation). For example, the number denoted by the expression

$$10 \uparrow\uparrow 10$$

is exceedingly big; most physically defined magnitudes (e.g., the ratio of the volume of the physical universe to the volume of an electron) are small compared to it. Nevertheless, it can easily be written in scientific notation, as follows. (Recall the parentheses convention.)

$$10\uparrow10\uparrow10\uparrow10\uparrow10\uparrow10\uparrow10\uparrow10\uparrow10\uparrow10$$

But the number denoted by the seemingly innocuous expression

$$10 \uparrow\uparrow\uparrow 10$$

cannot be expressed in ordinary scientific notation, not at least if we require that the final expression fit in the physical universe and that each constituent symbol be no smaller than an electron!

The numbers denoted by most hyper-exponential expressions are outrageously and unimaginably large, although they *are* finite. That they are finite does not mean they are trivial!

We conclude this section with examples of calculations involving hyper-exponentiation. First, a small one.

$$\begin{aligned}
2 \uparrow\uparrow\uparrow 3 &= 2 \uparrow\uparrow (2 \uparrow\uparrow 2) \\
&= 2 \uparrow\uparrow (2 \uparrow 2) \\
&= 2 \uparrow\uparrow 4 \\
&= 2 \uparrow (2 \uparrow (2 \uparrow 2))) \\
&= 2\uparrow16 \\
&= 65{,}536.
\end{aligned}$$

The next number we consider is:

$$\begin{aligned}
2 \uparrow\uparrow\uparrow 4 &= 2 \uparrow\uparrow (2 \uparrow\uparrow (2 \uparrow\uparrow 2)) \\
&= 2 \uparrow\uparrow (2 \uparrow\uparrow 4) \\
&= 2 \uparrow\uparrow 65{,}536 \\
&= \underbrace{2 \uparrow (2 \uparrow (2 \uparrow \ldots\ldots\ldots}_{\text{65,536 times!}}
\end{aligned}$$

This is a very large number, but it is easy to express in ordinary scientific notation; all we need is to write very small and use a very big sheet of paper! Still, it can be done.

The number $2\uparrow\uparrow\uparrow5$ is bigger still:

$$2 \uparrow\uparrow\uparrow 5 = \underbrace{2 \uparrow (2 \uparrow (2 \uparrow \ldots\ldots\ldots}_{2\uparrow\uparrow\uparrow4 \text{ times!}}$$

Now, we have completely outstripped scientific notation. And we haven't even discussed 4-th or 5-th order exponentiation. We conclude with an illustration of 5-th order exponentiation.

$2 \uparrow\uparrow\uparrow\uparrow\uparrow 3 = 2 \uparrow\uparrow\uparrow\uparrow (2 \uparrow\uparrow\uparrow\uparrow 2)$
$= 2 \uparrow\uparrow\uparrow\uparrow 4$
$= 2 \uparrow\uparrow\uparrow (2 \uparrow\uparrow\uparrow (2 \uparrow\uparrow\uparrow 2))$
$= 2 \uparrow\uparrow\uparrow (2 \uparrow\uparrow\uparrow 4)$
$= 2 \uparrow\uparrow (2 \uparrow\uparrow (2 \uparrow\uparrow .....$
+- - - - - - - - - - - - - - + ◄——— $2\uparrow\uparrow\uparrow\uparrow 3$ times!

11. The Algebra of Arithmetic

In the present section, we examine the algebraic properties of the standard arithmetic operations. We show that a few selected laws of arithmetic can be deduced from the postulates of set theory. Proof by induction is used extensively.

First, it is well-known that addition is both associative and commutative, which is to say that the following are theorems. [In this context, the domain of quantification is understood to be the set ω of natural numbers.]

(1) $\forall x \forall y \forall z[(x+y)+z = x+(y+z)]$ Ass(+)
(2) $\forall x \forall y[x+y = y+x]$ Com(+)

In order to prove (1), we do the outer two universal quantifiers by universal derivation, and we do the innermost universal quantifier by induction. We follow the following convention concerning constants: we use 'a', 'b', 'c' for constants used in ordinary universal derivation (UD), and we use 'm', 'n', 'k', 'j' for constants used in mathematical induction (MI).

	(1)	~~SHOW~~: $\forall x \forall y \forall z[(x+y)+z = x+(y+z)]$	UD2
	(2)	~~SHOW~~: $\forall z[(a+b)+z = a+(b+z)]$	MI
BC:	(3)	~~SHOW~~: $(a+b)+0 = a+(b+0)$	DD
	(4)	$(a+b)+0 = a+b$	Def +
	(5)	$b+0 = b$	Def +
	(6)	$a+(b+0) = a+b$	5,IL
	(7)	$(a+b)+0 = a+(b+0)$	4,6,IL
IC:	(*)	~~SHOW~~: $\forall x[(a+b)+x = a+(b+x)$	
		$\rightarrow (a+b)+x^+ = a+(b+x^+)$	UCD*
IH:	(8)	$(a+b)+m = a+(b+m)$	As
IS:	(9)	~~SHOW~~: $(a+b)+m^+ = a+(b+m^+)$	DD
	(10)	$(a+b)+m^+ = [(a+b)+m]^+$	Def +
	(11)	$= [a+(b+m)]^+$	IH
	(12)	$b+m^+ = (b+m)^+$	Def +
	(13)	$a+(b+m^+) = a+(b+m)^+$	12,IL
	(14)	$= [a+(b+m)]^+$	Def +
	(15)	$(a+b)+m^+ = a+(b+m^+)$	11,14,IL

Note that the line marked * is optional.

In the previous problem, in showing a triple-universal formula, we used universal derivation for two quantifiers, and induction for the third. We are not always so fortunate. The commutativity of addition, which involves two quantifiers, requires double-induction in its proof.

In order to see how double induction works, we first display ordinary induction schematically.

12. The Basic Scheme of Induction

	~~SHOW~~: $\forall x F[x]$	MI
BC:	~~SHOW~~: $F[0]$	??
IC:	~~SHOW~~: $\forall x(F[x] \ .\rightarrow \ F[x^+])$	UCD
IH:	$F[m]$	As
IS:	~~SHOW~~: $F[m^+]$	??

Note carefully here that the quantifiers/variables are understood to range over natural numbers. Now, the formula $F[x]$ can be any formula whatsoever; in particular, it can be a universal formula, in which case the method of induction is also appropriate for it. This yields the following scheme of double induction.

13. The Scheme of Double Induction

	~~SHOW~~: $\forall x \forall y F[x,y]$	MI
BC0:	~~SHOW~~: $\forall y F[0,y]$	MI
BC1:	~~SHOW~~: $F[0,0]$	??
IC1:	~~SHOW~~: $\forall y(F[0,y] \rightarrow F[0,y^+])$	UCD
IH1:	$F[0,m]$	As
IS1:	~~SHOW~~: $F[0,m^+]$	??
IC0:	~~SHOW~~: $\forall x(\forall y F[x,y] \rightarrow \forall y F[x^+,y])$	UCD
IH0:	$\forall y F[m,y]$	As
IS0:	~~SHOW~~: $\forall y F[m^+,y]$	MI
BC2:	~~SHOW~~: $F[m^+,0]$	??
IC2:	~~SHOW~~: $\forall y(F[m^+,y] \rightarrow F[m^+,y^+])$	UCD
IH2:	$F[m^+,n]$	As
IS2:	~~SHOW~~: $F[m^+,n^+]$	??

The basic idea is that an ordinary induction involves showing the base case and the inductive case. If these are both shown by induction, then each of these also involves both a base case and inductive case. In the above scheme, the outer cases are numbered 0, and the inner cases are numbered 1 and 2, respectively.

Before going on, we should note that the formula $F[x,y]$ may itself be a universally quantified formula, in which case all the lines marked by '??', of which there are four, may be shown by induction as well. As the reader may surmise, the general scheme of triple induction is even more complicated than the scheme of double induction.

The following proof of the commutativity of addition is an example of double induction.

	(1)	~~SHOW~~: $\forall x \forall y(x+y = y+x)$	MI
BC0:	(2)	~~SHOW~~: $\forall y(0+y = y+0)$	MI
BC1:	(3)	~~SHOW~~: $0+0 = 0+0$	IL
IC1:	(*)	~~SHOW~~: $\forall y(0+y=y+0 \rightarrow 0+y^+=y^++0$	UCD
IH1:	(4)	$0+m = m+0$	As
IS1:	(5)	~~SHOW~~: $0+m^+ = m^++0$	DD
	(6)	$0+m^+ = (0+m)^+$	Def +
	(7)	$0+m = m+0$	IH1
	(8)	$m+0 = m$	Def +
	(9)	$0+m^+ = m^+$	6,7,8, IL
	(10)	$m^++0 = m^+$	Def +
	(11)	$0+m^+ = m^++0$	9,10,IL
IC0:	(*)	~~SHOW~~: $\forall x[\forall y(m+y = y+m) \rightarrow \forall y(m^++y = y+m^+)]$	UCD
IH0:	(12)	$\forall y(m+y = y+m)$	As
IS0:	(13)	~~SHOW~~: $\forall y(m^++y = y+m^+)$	MI
BC2:	(14)	~~SHOW~~: $m^++0 = 0+m^+$	DD
	(15)	$m^++0 = m^+$	Def +
	(16)	$0+m^+ = (0+m)^+$	Def +
	(17)	$0+m = m+0$	IH0, IL
	(18)	$m+0 = m$	Def +
	(19)	$0+m = m$	17,18,IL
	(20)	$0+m^+ = m^+$	16,19,IL
	(21)	$m^++0 = 0+m^+$	15,20,IL
IC2:	(*)	~~SHOW~~: $\forall x(m^++x = n+m^+ \rightarrow m^++x^+ = x^++m^+)$	UCD
IH2:	(22)	$m^++n = n+m^+$	As
IS2:	(23)	~~SHOW~~: $m^++n^+ = n^++m^+$	DD
	(24)	$m^++n^+ = (m^++n)^+$	Def +
	(25)	$= (n+m^+)^+$	24,IH2
	(26)	$= (n+m)^{++}$	Def +
	(27)	$n^++m^+ = (n^++m)^+$	Def +
	(28)	$= (m+n^+)^+$	IH0 (12)
	(29)	$= (m+n)^{++}$	Def +
	(30)	$= (n+m)^{++}$	IH0
	(31)	$m^++n = n+m^+$	26,30,IL

We have dealt exclusively with addition. We conclude this section with a proof of a well-known law about the inter-action between multiplication and addition. In this proof, we assume the previous principles about addition – the commutativity and associativity of addition. For the sake of visual simplicity, we write 'ab' in place of '$a \times b$', as is customary.

	(1)	~~SHOW~~: $\forall x \forall y \forall z[x(y+z) = xy+xz]$	UD2
	(2)	~~SHOW~~: $\forall z[a(b+z) = ab+az]$	MI
BC:	(3)	~~SHOW~~: $a(b+0) = ab+a0$	DD
	(4)	$b+0 = b$	Def +
	(5)	$a(b+0) = ab$	4,IL
	(6)	$a0 = 0$	Def ×
	(7)	$ab+a0 = ab+0$	6,IL
	(8)	$= ab$	Def +
	(9)	$a(b+0) = ab+a0$	5,8,IL
IC:	(*)	~~SHOW~~: $\forall x[a(b+x) = ab+ax \rightarrow a(b+x+) = ab + ax+]$	UCD
IH:	(10)	$a(b+m) = ab+am$	As
IS:	(11)	~~SHOW~~: $a(b+m+) = ab+am+$	DD
	(12)	$b+m+ = (b+m)+$	Def +
	(13)	$a(b+m+) = a(b+m)+$	12,IL
	(14)	$= a+(a(b+m))$	Def ×
	(15)	$= a+(ab+am)$	IH,IL
	(16)	$am+ = a+am$	Def ×
	(17)	$ab+am+ = ab+(a+am)$	16,IL
	(18)	$= (ab+a)+am$	Ass(+)
	(19)	$= (a+ab)+am$	Com(+),IL
	(20)	$= a+(ab+am)$	Ass(+)
	(21)	$a(b+m+) = ab+am+$	15,20,IL

Parenthesis convention:

$(a \times b)+c$ is written $ab+c$
$a \times (b+c)$ is written $a(b+c)$
$a+b^+$ is a plus the successor of b
$(a+b)^+$ is the successor of $a+b$

14. The Order-Theory of Arithmetic

An important remaining feature of the theory of natural numbers is the order-theoretic structure of ω, which pertains to the less-than and less-than-or-equal-to relations.

Our definition of the natural numbers makes the definition of the less-than predicate (written <) very easy:

(D10) $a < b \;=_{df}\; a \in b$

The corresponding less-than-or-equal-to predicate is then defined in the usual way.

(D11) $a \leqslant b \;=_{df}\; a < b \vee a = b$

We have defined the relevant *predicates*. The associated *relations* (sets of ordered pairs) are defined as follows.

(d1) $<_\omega \;=_{df}\; \{ (x,y): x,y \in \omega \;\&\; x < y \}$

(d2) $\leqslant_\omega \;=_{df}\; \{ (x,y): x,y \in \omega \;\&\; x \leqslant y \}$

It is customary to abuse notation slightly, and use the same symbol for both the predicate and the relation, so long as the field of the relation is clear. Since we are talking exclusively about natural numbers, We follow this custom, and drop the subscript 'ω'.

Definitions are cheap, theorems less so. Having defined the order-theoretic predicates/relations for arithmetic, we still need to prove the relevant theorems. Some of the more prominent theorems are the following (note that '<' abbreviates '$<_\omega$'.)

(t1) $<$ is transitive, asymmetric, irreflexive, and weakly connected.

(t2) $\leqslant$ is transitive, anti-symmetric, reflexive, and strongly connected.

(t3) $\leqslant$ is a well-ordering (on ω).

[See later chapter concerning well-ordering.]

A corollary to (t1) is the Law of Trichotomy, written as follows.

(t4) for any pair of numbers m,n, exactly one of the following obtains:

(1) $m < n$
(2) $n < m$
(3) $m = n$

Let us examine Theorem (t1), which says that $\mathcal{P}$ is transitive, asymmetric, irreflexive, and weakly connected, which is to say the following, where the quantifiers range over ω.

(tr) $\forall x \forall y \forall z(x < y \ \& \ y < z \ .\rightarrow \ x < z)$
(asy) $\forall x \forall y(x < y \ \rightarrow \ \sim[y < x])$
(irr) $\forall x \sim[x < x]$
(wc) $\forall x \forall y(x \neq y \ \rightarrow . \ x < y \ \vee \ y < x)$

Given the definition of '<' in terms of '$\in$', the transitivity of $\mathcal{P}$ is an immediate consequence of the theorem that every natural number is a transitive set. Similarly, the asymmetry of $<$ is an immediate consequence of the theorem that no number is an element of any of its elements. The irreflexivity of $<$ is a logical consequence of its transitivity and asymmetry.

Proving the weak connectivity of $<$ is more difficult. First, we formally write down the transitivity of $<$ as Lemma 0, for later reference. Next, we write down and sketch the proofs of three further lemmas.

Lemma 0: $\forall x \forall y \forall z(x < y \ \& \ y < z \ .\rightarrow \ x < z)$

Lemma 1: $\forall x[x < x^+]$

Proof: $a \in \{a\}$, so $a \in a \cup \{a\}$, so $a \in a^+$, so $a < a^+$.

Lemma 2: $\forall x(x \in \omega \ \rightarrow . \ 0 < x \ \vee \ 0 = x)$

Proof (by induction): BC: trivial. IH: assume $0 < m \ \vee \ 0 = m$, to show (IS): $0 < m^+ \ \vee \ 0 = m^+$. We can in fact show: $0 < m^+$. Now, $m = 0 \vee m \neq 0$. Case 1: $m = 0$, in which case $m^+ = \{0\}$. But $0 \in \{0\}$, so $0 \in m^+$, and hence $0 < m^+$. Case 2: $m \neq 0$, so by (IH), we have $0 < m$, but by lemma 1, $m < m^+$, so by lemma 0, $0 < m^+$.

Lemma 3: $\forall x \forall y(x < y \rightarrow x^+ < y^+)$

Proof: UD (MI): BC: $\forall y(a < 0 \rightarrow a^+ < a)$ (trivial since antecedent is impossible). IH: $a < n \rightarrow a^+ < n^+$, to show IS: $a < n^+ \rightarrow a^+ < n^{++}$. Assume: $a < n^+$, to show: $a^+ < n^{++}$. By definition of $<$, $a \in n^+$, which means that $a \in n \cup \{n\}$. Case 1: $a \in n$, in which case $a < n$, so by IH, $a^+ < n^+$, and hence, by lemmas 0,1, $a^+ < n^{++}$ (qed). Case 2: $a = n$, in which case $a^+ = n^+$. But $n^+ < n^{++}$, by lemma 1, so by identity logic, $a^+ < n^{++}$ (qed).

Now, we provide a more formal proof of (wc), for which we employ (double) induction, together with the above lemmas and previous theorems. Notice that Base Case #0 is not shown by induction, so there is no need for BC1, IC1, IH1, IS1. Also notice that the inductive cases are written purely schematically as 'show universal conditional.

	(1)	~~SHOW~~: $\forall x \forall y(x \neq y \rightarrow . \; x < y \vee y < x)$	MI
BC0:	(2)	~~SHOW~~: $\forall y(0 \neq y \rightarrow . \; 0 < y \vee y < 0)$	
	(3)	follows from lemma 2.	
IC0:	(*)	~~SHOW~~: UC	UCD
IH0:	(4)	$\forall y(m \neq y \rightarrow . \; m < y \vee y < m)$	As
IS0:	(5)	~~SHOW~~: $\forall y(m^+ \neq y \rightarrow . \; m^+ < y \vee y < m^+)$	MI
BC2:	(6)	~~SHOW~~: $m^+ \neq 0 \rightarrow . \; m^+ < 0 \vee 0 < m^+$	
	(7)	follows from lemma 2.	
IC2:	(*)	~~SHOW~~: UC	UCD
IH2:	(8)	$m^+ \neq n \rightarrow . \; m^+ < n \vee n < m^+$	As
IS2:	(9)	~~SHOW~~: $m^+ \neq n^+ \rightarrow . \; m^+ < n^+ \vee n^+ < m^+$	CD
	(10)	$m^+ \neq n^+$	As
	(11)	~~SHOW~~: $m^+ < n^+ \vee n^+ < m^+$	SC
	(12)	$m \neq n$	10, PT
	(13)	$m < n \vee n < m$	4,12,QL
	(14)	c1: $m < n$	As
	(15)	$m+ < n+$	14, lemma 3
	(16)	c2: $n < m$	As
	(17)	$n+ < m+$	16, lemma 3

The following are theorems concerning the interaction between the order-structure and the algebraic structure of the numbers.

(t5) $\forall x \forall y \forall z(x < y \rightarrow x+z < y+z)$
(t6) $\forall x \forall y \forall z(z \neq 0 \rightarrow (x < y \rightarrow xz < yz))$
(t7) $\forall x \forall y \forall z(x < y \rightarrow xz^+ < yz^+)$
(t8) $\forall x \forall y(x < y \leftrightarrow \exists z(z \neq 0 \;\&\; x+z = y))$

For the sake of illustration, we prove (t5), which involves a single induction (over z).

	(1)	~~SHOW~~: $\forall x \forall y \forall z(x < y \rightarrow x+z < y+z)$	UD2
	(2)	~~SHOW~~: $\forall z(a < b \rightarrow a+z < b+z)$	MI
BC:	(3)	~~SHOW~~: $a < b \rightarrow a+0 < b+0$	CD
	(4)	$a < b$	As
	(5)	~~SHOW~~: $a+0 < b+0$	DD
	(6)	$a+0=a$	def +
	(7)	$b+0=b$	def +
	(8)	$a+0 < b+0$	4,6,7,IL
IC:	(*)	~~SHOW~~: UC	UCD
IH:	(9)	$a < b \rightarrow a+m < b+m$	As
IS:	(10)	~~SHOW~~: $a < b \rightarrow a+m^+ < b+m^+$	CD
	(11)	$a < b$	As
	(12)	~~SHOW~~: $a+m^+ < b+m^+$	DD
	(13)	$a+m < b+m$	9,11,SL
	(14)	$(a+m)^+ < (b+m)^+$	13,lemma 3
	(15)	$a+m^+ < b+m^+$	14, def +

15. Definitions for Chapter 4

1. $0xF[x] =_{df} \sim\exists xF[x]$
   $1xF[x] =_{df} \exists x\forall y(F[y] \leftrightarrow y{=}x)$
   $2xF[x] =_{df} \exists x\exists y(x{\neq}y \;\&\; \forall z(F[z] \leftrightarrow .\; z{=}x \vee z{=}y))$
   etc.

2. $0[A] =_{df} 0x[x \in A]$
   $1[A] =_{df} 1x[x \in A]$
   $2[A] =_{df} 2x[x \in A]$
   etc.

3. $0[A] =_{df} \sim\exists x[x \in A]$
   $1[A] =_{df} \exists x\forall y(y \in A \;\leftrightarrow\; y{=}x)$
   $2[A] =_{df} \exists x\exists y(x{\neq}y \;\&\; \forall z(z \in A \;\leftrightarrow .\; z{=}x \vee z{=}y))$
   etc.

4. $0 =_{df} \varnothing$
   $1 =_{df} \{0\}$
   $2 =_{df} \{0,1\}$
   etc.

5. $A^{+} =_{df} A \cup \{A\}$

6. $suc[A] =_{df} 0 \in A \;\&\; \forall x(x \in A \;\rightarrow\; x^{+} \in A)$

7. A is the smallest $F[X] =_{df} F[A] \;\&\; \forall X(F[X] \rightarrow A \subseteq X)$.

8. $\omega =_{df}$ the smallest successor set

9. $fun(A,a,g) =_{df} \cap\{R : R \subseteq \omega \times A \;\&\; 0Ra \;\&\; \forall x\forall y[xRy \;\rightarrow\; x^{+}Rg(y)]\}$

10. $m{+}0 = m$
    $m{+}n^{+} = (m{+}n)^{+}$

11. $m{\times}0 = 0$
    $m{\times}n^{+} = m{+}(m{\times}n)$

12. $m{\uparrow}0 = 1$
    $m{\uparrow}n^{+} = m{\times}(m{\uparrow}n)$

13. $m{\uparrow\uparrow}0 = 1$
    $m{\uparrow\uparrow}n^{+} = m{\uparrow}(m{\uparrow\uparrow}n)$

14. $m{\uparrow\uparrow\uparrow}0 = 1$
    $m{\uparrow\uparrow\uparrow}n^{+} = m{\uparrow\uparrow}(m{\uparrow\uparrow\uparrow}n)$

15. $a < b =_{df} a \in b$

16. $a \leqslant b =_{df} a < b \vee a{=}b$

16. Theorems for Chapter 4

(1) $\forall x(x \in \omega \rightarrow \text{trans}[x])$

(2) $trans[\omega]$

(3) $\forall x(x \in \omega \rightarrow \forall y(y \in x \rightarrow x \notin y))$

(4) $\forall x(x \in \omega \rightarrow \forall y(y \in x \rightarrow \sim[x \subseteq y]))$

(5) $\sim\exists x[x^+ = 0]$

(6) $\forall x(x \in \omega \rightarrow x \neq x^+)$

(7) $\forall x(x \in \omega \rightarrow x \notin x)$

(8) $\forall x(x \in \omega \rightarrow (x \neq 0 \rightarrow \exists y(y \in \omega \ \& \ x = y^+)))$

(9) $\forall x \forall y(x,y \in \omega \rightarrow .\ x^+ = y^+ \rightarrow x = y)$

(10a) $\forall x(x \in \omega \rightarrow \bigcup x \subseteq x)$

(10b) $\bigcup\omega = \omega$

In each of the following, we presume that the quantifiers range over natural numbers (so '$\forall x\mathbb{F}$' abbreviates '$\forall x(x \in \omega \rightarrow \mathbb{F})$', and '$\exists x\mathbb{F}$' abbreviates '$\exists x(x \in \omega \ \& \ \mathbb{F})$'. In this context, '$ab$' is short for '$a \times b$'.

(11) $\forall x \forall y \forall z[x+(y+z) = (x+y)+z]$

(12) $\forall x \forall y[x+y = y+z]$

(13) $\forall x \forall y(x+y = x \rightarrow y = 0)$

(14) $\forall x \forall y \forall z[x(y+z) = xy+xz]$

(15) $\forall x \forall y \forall z[x(yz) = (xy)z]$

(16) $\forall x \forall y[xy = yx]$

(17) $\forall x(\forall y[xy = y] \rightarrow x = 1)$

(18) $\sim\forall x \forall y \forall z[x \uparrow (y \uparrow z) = (x \uparrow y) \uparrow z]$

(19) $\sim\forall x \forall y[x \uparrow y = y \uparrow x]$

(20) $\forall x \forall y \forall z[x \uparrow (y+z) = (x \uparrow y)(x \uparrow z)]$

(21a) $tr[<]$

(21b) $asy[<]$

(21c) $irr[<]$

(21d) $wcon[<]$

(22a) $tr[\leqslant]$

(22b) $ant[\leqslant]$

(22c) $ref[\leqslant]$

(22d) $scon[\leqslant]$

(23) $\forall x \forall y \ \text{XOR}(x < y, y < x, x = y)$

(24) $\forall x \forall y(x < y \leftrightarrow x \subset y)$

(25) $\forall x \forall y(x \leqslant y \leftrightarrow x \subseteq y)$

(26) $\forall x \forall y(x \leqslant y \leftrightarrow x \in y^+)$

(27) $\forall x \forall y \forall z(x < y \rightarrow x+z < y+z)$

(28) $\forall x \forall y \forall z(z \neq 0 \rightarrow (x < y \rightarrow xz < yz))$

(29) $\forall x \forall y \forall z(x < y \rightarrow xz^+ < yz^+)$

(30) $\forall x \forall y(x < y \leftrightarrow \exists z(z \neq 0 \ \& \ x+z = y))$

17. Summary of Mathematical Induction

1. General Principle of Mathematical Induction

(PMI) $F[0] \,\&\, \forall x(x \in \omega \,\&\, F[x] \mathrel{.\!\!\rightarrow} F[x^+]) \mathrel{.\!\!\rightarrow} \forall x(x \in \omega \rightarrow F[x])$

2. Special Principle of Mathematical Induction

[where quantifiers are presumed to range over numbers]

(SPMI) $F[0] \,\&\, \forall x(\, F[x] \rightarrow F[x^+]\,) \mathrel{.\!\!\rightarrow} \forall x F[x]$

3. The Basic Scheme for Proof by Induction

(The following presumes that the quantifiers range over numbers.)

	~~SHOW~~: $\forall x F[x]$	MI
BC:	~~SHOW~~: $F[0]$	??
IC:	~~SHOW~~: $\forall x(F[x] \mathrel{.\!\!\rightarrow} F[x^+])$	UCD *
IH:	$F[m]$	As
IS:	~~SHOW~~: $F[m^+]$	??

4. The Scheme of Double Induction

(which presumes quantifiers range over numbers)

	~~SHOW~~: $\forall x \forall y F[x,y]$	MI
BC0:	~~SHOW~~: $\forall y F[0,y]$	MI
BC1:	~~SHOW~~: $F[0,0]$	??
IC1:	~~SHOW~~: $\forall y(F[0,y] \rightarrow F[0,y+])$	UCD *
IH1:	$F[0,m]$	As
IS1:	~~SHOW~~: $F[0,m+]$	??
IC0:	~~SHOW~~: $\forall x(\forall \mathrm{yF}[x,y] \rightarrow \forall \mathrm{yF}[x^+,y])$	UCD *
IH0:	$\forall \mathrm{yF}[m,y]$	As
IS0:	~~SHOW~~: $\forall \mathrm{yF}[m^+,y]$	MI
BC2:	~~SHOW~~: $F[m^+,0]$	??
IC2:	~~SHOW~~: $\forall y(F[m^+,y] \rightarrow F[m^+,y^+])$	UCD *
IH2:	$F[m^+,n]$	As
IS2:	~~SHOW~~: $F[m^+,n^+]$	??

(Lines marked with '*' are optional in proof.)

18. Examples of Derivations for Chapter 4

See end of second page for lemmas employed in these derivations. Notice that 'BC:' and 'IS:' are short for '~~SHOW~~:', and that the inductive case lines and their boxes are omitted. ILD is a combination of IL and DD; basically, to show '$a=b$' we show '$a=c$' and '$b=c$'.

#15

(1)	~~SHOW~~: $\forall x \forall y \forall z[x(yz) = (xy)z]$	UD2
(2)	~~SHOW~~: $\forall z[a(bz) = (ab)z]$	MI
(3)	BC: $a(b0) = (ab)0$	ILD
(4)	$a(b0) = a0$	def $\times$
(5)	$= 0$	def $\times$
(6)	$(ab)0 = 0$	def $\times$
(7)	IH: $a(bm) = (ab)m$	As
(8)	IS: $a(bm+) = (ab)m+$	ILD
(9)	$a(bm+) = a(b+bm)$	def $\times$
(10)	$= ab + a(bm)$	dist
(11)	$= ab + (ab)m$	IH (7)
(12)	$(ab)m+ = ab + (ab)m$	def $\times$

#16

(1)	~~SHOW~~: $\forall x \forall y[xy = yx]$	MI
(2)	BC: $\forall y[0y = y0]$	MI
(3)	BC: $00 = 00$	IL
(4)	IH: $0m = m0$	As
(5)	IS: $0m^+ = m^+0$	DD (IL)
(6)	$0m^+ = 0+0m$	def $\times$
(7)	$= 0+m0$	IH (4)
(8)	$= 0+0$	def $\times$
(9)	$= 0$	def +
(10)	$m^+0 = 0$	def $\times$
(11)	IH: $\forall y[my = ym]$	As
(12)	IS: $\forall y[m^+y = ym^+]$	MI
(13)	BC: $m^+0 = 0m^+$	2
(14)	IH: $m^+n = nm^+$	As
(15)	IS: $m^+n^+ = n^+m^+$	DD (IL)
(16)	$m^+n^+ = m^+ + m^+n$	def $\times$
(17)	$= m^+ + nm^+$	IH(14)
(18)	$= m^+ + (n+nm)$	def $\times$
(19)	$= (n+nm) + m^+$	com[+]
(20)	$= [(n+nm)+m]^+$	def +
(21)	$n^+m^+ = n^+ + n^+m$	def $\times$
(22)	$= n^+ + mn^+$	IH(11)
(23)	$= n^+ + (m + mn)$	def $\times$
(24)	$= (m+mn) + n^+$	com[+]
(25)	$= [(m+mn)+n]^+$	def +
(26)	$= [(m+nm)+n]^+$	IH(11)
(27)	$= [n+(m+nm)]^+$	ass[+]
(28)	$= [n+(nm+m)]^+$	com[+]
(29)	$= [(n+nm)+m]^+$	ass[+]

#28

(1)	~~SHOW~~: $\forall x \forall y(x < y \rightarrow \forall z(z \neq 0 \rightarrow xz < yz))$	UCD
(2)	$\quad a < b$	As
(3)	$\quad$ ~~SHOW~~: $\forall z(z \neq 0 \rightarrow az < bz))$	MI
(4)	$\quad\quad$ BC: $0 \neq 0 \rightarrow a0 < b0$	IL
(5)	$\quad\quad$ IH: $m \neq 0 \rightarrow am < bm$	As
(6)	$\quad\quad$ IS: $m^+ \neq 0 \rightarrow am^+ < bm^+$	CD
(7)	$\quad\quad\quad m^+ \neq 0$	As
(8)	$\quad\quad\quad$ ~~SHOW~~: $am^+ < bm^+$	SC
(9)	$\quad\quad\quad$ c1: $m = 0$	As
(10)	$\quad\quad\quad am^+ \quad = a0^+$	IL
(11)	$\quad\quad\quad\quad\quad = a+a0$	Def $\times$
(12)	$\quad\quad\quad\quad\quad = a+0$	Def $\times$
(13)	$\quad\quad\quad\quad\quad = a$	Def +
(14)	$\quad\quad\quad bm^+ = b$	similarly
(15)	$\quad\quad\quad am^+ < bm^+$ (qed)	2,13,14,IL
(16)	$\quad\quad\quad$ c2: $m \neq 0$	As
(17)	$\quad\quad\quad am < bm$	IH
(18)	$\quad\quad\quad a+am < a+bm$	lemma 5a
(19)	$\quad\quad\quad a+bm < b+bm$	lemma 5b
(20)	$\quad\quad\quad a+am < b+bm$	lemma 4
(21)	$\quad\quad\quad am^+ = a+am$	Def $\times$
(22)	$\quad\quad\quad bm^+ = b+bm$	Def $\times$
(23)	$\quad\quad\quad am^+ < bm^+$ (qed)	20,21,22,IL

Lemma 1: $a+(b+c) = (a+b)+c$	ass[+]
Lemma 2: $a+b = b+a$	com[+]
Lemma 3: $a(b+c) = ab+ac$	dist
Lemma 4: $a < b \ \& \ b < c \ .\rightarrow \ a < c$	tr[<]
Lemma 5a: $a < b \ \rightarrow \ a+c < b+c$	
Lemma 5b: $a < b \ \rightarrow \ c+a < c+b$	
Lemma 6: $a \neq 0 \ \rightarrow \ \exists x[a = x^+]$	
Lemma 7: $\sim[a < 0]$	
Lemma 8: $a < b^+ \leftrightarrow . \ a < b \ \vee \ a = b$	
Lemma 9: $a^+ = b^+ \ \rightarrow \ a = b$	
Lemma 10: $a < a^+$	
Lemma 11: $\sim\exists x[x^+ = 0]$	

#30

(1)	~~SHOW~~: $\forall x \forall y(x < y \leftrightarrow \exists z(z \neq 0 \ \& \ x+z=y))$	UD
(2)	~~SHOW~~: $\forall y(a < y \leftrightarrow \exists z(z \neq 0 \ \& \ a+z=y))$	MI
(3)	BC: $a < 0 \leftrightarrow \exists z(z \neq 0 \ \& \ a+z=0)$	DD (SL)
(4)	$\sim[a < 0]$	lemma 7
(5)	~~SHOW~~: $\sim\exists z(z \neq 0 \ \& \ a+z=0)$	ID
(6)	$\exists z(z \neq 0 \ \& \ a+z=0)$	As
(7)	~~SHOW~~: $\times$	DD
(8)	$b \neq 0 \ \& \ a+b=0$	6,$\exists$O
(9)	$\exists x[b=x^+]$	8,lemma 6
(10)	$b=c^+$	9,$\exists$O
(11)	$a+c^+=0$	8,10
(12)	$(a+c)^+=0$	ef +
(13)	$\exists x[x^+=0]$	12,QL
(14)	$\sim\exists x[x^+=0]$	lemma 11
(15)	IH: $a < m \leftrightarrow \exists z(z \neq 0 \ \& \ a+z = m)$	As
(16)	IS: $a < m^+ \leftrightarrow \exists z(z \neq 0 \ \& \ a+z = m^+)$	$\leftrightarrow$D
(17)	$a < m^+$	As
(18)	~~SHOW~~: $\exists z(z \neq 0 \ \& \ a+z = m^+)$	SC
(19)	$a < m \ \vee \ a=m$	17, lemma 8
(20)	c1: $a < m$	As
(21)	$\exists z(z \neq 0 \ \& \ a+z = m)$	15,20
(22)	$b \neq 0 \ \& \ a+b = m$	21,$\exists$O
(23)	$b^+ \neq 0$	lemma 11
(24)	$a+b^+ = (a+b)^+$	def +
(25)	$a+b^+ = m^+$	22,24
(26)	$b^+ \neq 0 \ \& \ a+b^+ = m^+$	23,25
(27)	$\exists z(z \neq 0 \ \& \ a+z = m^+)$	26,QL
(28)	c2: $a=m$	As
(29)	$a+0^+ = m+0^+$	IL
(30)	$= (m+0)^+$	def +
(31)	$= m^+$	def +
(32)	$0^+ \neq 0$	lemma 11
(33)	$\exists z(z \neq 0 \ \& \ a+z = m^+)$	31,32,QL
(34)	$\exists z(z \neq 0 \ \& \ a+z = m^+)$	As
(35)	~~SHOW~~: $a < m^+$	DD
(36)	$b \neq 0 \ \& \ a+b = m^+$	34,$\exists$O
(37)	$\exists x[b=x^+]$	36,lemma 6
(38)	$b=c^+$	37,$\exists$O
(39)	$m^+ = a+b$	36,IL
(40)	$= a+c^+$	38,39,IL
(41)	$= (a+c)^+$	def +
(42)	$a+b = m$	41,lemma 9
(43)	$a < m$	36,42,IH
(44)	$m < m^+$	lemma 10
(45)	$a < m^+$	43,44, tr[<]