CHAPTER 2: THE NATURAL NUMBERS $\mathbb{N}$ AS AN ORDERED SET.

WAYNE AITKEN AND LINDA HOLT

Summer 2019 Edition

In this chapter we continue the study of $\mathbb{N}$ with a focus on the definition of order $<$ and matters related to this order. For instance we will show that $\mathbb{N}$ is well-ordered. We will define subtraction in a manner related to the order. We will discuss the set $\{1, \ldots, n\}$ which requires the order on $\mathbb{N}$ to define. Finally, we will discuss recursion which, in a sense, also depends on order since since a value of a function $f(n)$ is defined in terms of $f(m)$ for one or more $m < n$.

We start with the general concept of a strict linear order. Then we consider the specific order relation on $\mathbb{N}$, and then treat various topics related to this order.

1. Order relations

Several of the number systems considered in this course can be thought of as having a linear arrangement. Such number systems include $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, and $\mathbb{R}$. In this section we defined the general notion of a linear order. In the next section we define the specific linear order for $\mathbb{N}$.

Definition 1. A *strict linear order* on a set S is a relation, commonly written with the symbol $<$, which satisfies the following two conditions.

(1) *The transitivity law*: for all $x, y, z \in S$,

$$\text{if } x < y \text{ and } y < z \text{ then } x < z.$$

(2) *The trichotomy law*: for all $x, y \in S$ exactly one of the following three holds

$$x < y, \qquad x = y, \qquad y < x.$$

Remark 1. When $x < y$ holds we say that x *is strictly less than* y. Strict linear orders are also called *strict total orders*, in contrast to a more general type of order relation important in mathematics called *partial orders*.

Remark 2. (Review) Order relations are one of the two types of binary relations used in this course. The other type are the *equivalence relations*.

Equivalence relations are relations which are reflexive, symmetric, and transitive.

We assume that that reader is familiar with the general idea of a binary relation. Informally, a binary relation R on a set S is a condition linking elements of S. We write xRy to indicate that $x, y \in S$ are linked according to the relation R, and $\neg xRy$ to indicate otherwise. For each pair $(x, y) \in S \times S$ we either have xRy or $\neg\, xRy$.

In modern set theory, it is usual to think of a relation R on S as a subset of $S \times S$. For such a relation $R \subseteq S \times S$ we have $(x, y) \in R$ if and only if the relation holds between x and y. In other words, xRy just means that $(x, y) \in R$. If $(x, y) \notin R$ then we write $\neg\, xRy$ instead.

The convention of putting the symbol R between the elements x, y is called *infix notation*. We write xRy when using infix notation, versus writing $R(x, y)$ or $R\, x\, y$ using prefix notation. For equivalence relations, we often use symbols such as $\equiv$ or $\sim$ for the relation R. For strict linear orders the common symbol for the relation R is $<$, as in the definition above.

Exercise 1. Suppose that $<$ is a strict linear order on a set S. Show that this relation is *anti-reflexive*. In other words, show that $\neg(x < x)$ for all $x \in S$.

Exercise 2. There are other equivalent ways of defining linear orders. For example, instead of trichotomy, one could require "weak trichotomy" together with the anti-reflexive law. With this in mind, prove the following lemma.

Lemma 1. *Suppose $<$ is a binary relation on a set S such that the following three conditions hold:*

(1) *The transitivity law: for all $x, y, z \in S$,*
$$\text{if } x < y \text{ and } y < z \text{ then } x < z.$$
(2) *The weak trichotomy law: for all $x, y \in S$*
$$(x < y) \vee (x = y) \vee (y < x).$$
(3) *The anti-reflexive law: for all $x \in S$*
$$\neg(x < x).$$

Given these all hold, then $<$ is a strict linear order on S.

Informal Exercise 3. What is the main difference between the trichotomy law and the weak trichotomy law? Which implies the other?

Definition 2. Let $<$ be a strict linear order on a set S. Then we define the associated nonstrict order relation $\leq$ as follows: For each $x, y \in S$ the relation $x \leq y$ is defined to hold if and only if
$$(x < y) \ \vee \ (x = y)$$
holds. When $x \leq y$ holds we say that x *is less than or equal to* y.

Exercise 4. Let $\leq$ be as in the above definition. Show the following:

(1) $(x \leq y) \vee (y \leq x)$ for all $x, y \in S$.
(2) If $x \leq y$ and $y \leq x$ then $x = y$.
(3) $\neg(x \leq y)$ if and only if $y < x$.
(4) $\neg(x < y)$ if and only if $y \leq x$.

Exercise 5. From the transitivity of $<$ we can derive other transitivity laws. Do this by proving the following two theorems.

Theorem 2 (mixed transitivity)**.** *Suppose that $<$ is a strict linear order on a set S. Let $\leq$ be as in Definition 2. Then*
(i) For all $x, y, z \in S$, if $x < y$ and $y \leq z$ then $x < z$.
(ii) For all $x, y, z \in S$, if $x \leq y$ and $y < z$ then $x < z$.

Theorem 3 (transitivity of $\leq$)**.** *Suppose that $x, y, z \in S$ and that $<$ is a strict linear order on a set S. Let $\leq$ be as in Definition 2. If $x \leq y$ and $y \leq z$ then $x \leq z$.*

Definition 3. Suppose that $<$ is a strict linear order on a set S. Then we define $>$ as follows: for all $x, y \in S$ the relation $x > y$ holds if and only if $y < x$ holds. In this case we say that x *is strictly greater than* y.

We define $\geq$ as follows: for all $x, y \in S$ the relation $x \geq y$ holds if and only if $(x > y) \vee (x = y)$ holds. In this case we say that x *is greater than or equal to* y.

Remark 3. Since $<$ and $\leq$ are transitive, it follows easily that $>$ and $\geq$ are transitive as well.

Definition 4. The notation $a < b < c$ is short for $(a < b) \wedge (b < c)$. Observe that by transitivity $a < b < c$ implies $a < c$.

A similar notation is adopted for $>$, $\leq$, and $\geq$. We adopt a similar notation also for more than three terms, and we sometimes combine $<$ and $\leq$ (or $>$ and $\geq$). Thus $a \leq b < c < d$ is short for $(a \leq b) \wedge (b < c) \wedge (c < d)$.

2. The standard order relations on $\mathbb{N}$

Our definition of order for $\mathbb{N}$ is based on addition:

Definition 5. We define the binary relation $<$ on $\mathbb{N}$ by the following rule: Given $m, n \in \mathbb{N}$, we have $m < n$ if and only if there is a nonzero $b \in \mathbb{N}$ such that $n = m + b$. In symbols:

$$m < n \iff \exists b \in \mathbb{N} \big((b \neq 0) \wedge (n = m + b) \big).$$

We get two quick consequences of this definition, which we state for future reference.

Theorem 4. *If $x, y \in \mathbb{N}$ and if $y \neq 0$ then $x < x + y$.*

Theorem 5. *If $m \in \mathbb{N}$ then $m < m + 1$.*

Exercise 6. Give short proofs of the above two theorems. Hint: do we know that $1 \neq 0$?

Theorem 6 (transitivity of $<$)**.** *Suppose $x, y, z \in \mathbb{N}$. If $x < y$ and $y < z$ then $x < z$.*

Proof. Since $x < y$, there is a $b \neq 0$ such that $y = x + b$, and since $y < z$ there is a $c \neq 0$ such that $z = y + c$. So

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

In Chapter 1 we proved that $b + c \neq 0$. Therefore, $x < z$ by Definition 5 □

Lemma 7. *If $n \in \mathbb{N}$ then $(0 < n) \vee (0 = n)$.*

Proof. If $n = 0$ then the result holds. So assume $n \neq 0$. In this case we have $0 < 0 + n$ by Theorem 4. Then since $0 + n = n$, we get $0 < n$. □

We now give the proof of the weak trichotomy law. Its proof is a fairly tricky use of induction.

Lemma 8 (Weak trichotomy)**.** *Suppose $m, n \in \mathbb{N}$. Then*

$$(m < n) \vee (m = n) \vee (n < m).$$

Proof. Let $n \in \mathbb{N}$ be fixed. Let S_n be the set of elements $x \in \mathbb{N}$ that satisfy the following condition:

$$(x < n) \vee (x = n) \vee (n < x).$$

In other words, S_n is the set of all m for which the lemma holds (with fixed n).

The base case $0 \in S_n$ follows from Lemma 7.

Now suppose that $k \in S_n$. We wish to show $\sigma k \in S_n$. Since $k \in S_n$ we have three cases: (1) $k < n$, (2) $k = n$, and (3) $n < k$.

Case 1: $k < n$. By definition $n = k + b$ for some $b \neq 0$. Since b is not 0, it has a predecessor $c \in \mathbb{N}$. So $b = c + 1$ and

$$n = k + b = k + (c + 1) = k + (1 + c) = (k + 1) + c = \sigma k + c.$$

If $c = 0$ then $\sigma k = n$, so $\sigma k \in S_n$. If $c \neq 0$ then $\sigma k < n$ by Definition 5. so $\sigma k \in S_n$. So whatever c is, we have $\sigma k \in S_n$.

Case 2: $k = n$. Observe that $k < k + 1$ by Theorem 5. So $n < \sigma k$ since $k = n$ and $\sigma k = k + 1$. So $\sigma k \in S_n$.

Case 3: $n < k$. Observe that $k < k + 1$ by Theorem 5. So $n < k + 1$ by the transitivity of $<$. In other words $n < \sigma k$. Thus $\sigma k \in S_n$.

By the induction axiom, $S_n = \mathbb{N}$. This is true of arbitrary $n \in \mathbb{N}$. So for any $m, n \in \mathbb{N}$, we have $m \in S_n$. The result follows. □

Exercise 7. Use the cancellation law for addition (or related properties) to show the following: if $n \in \mathbb{N}$ then $\neg(n < n)$. Now use this fact, together with transitivity and weak trichotomy to conclude the following:

Theorem 9. *The relation $<$ is a strict linear order on $\mathbb{N}$.*

Remark 4. We define $>$, $\leq$, and $\geq$ in terms of $<$ as in the previous section.

We can now show that 0 is the least element of $\mathbb{N}$:

Theorem 10. *If $n \in \mathbb{N}$ then $0 \leq n$.*

Proof. This follows from the definition of $\leq$ and Lemma 7. □

Theorem 11. *Suppose $m, n \in \mathbb{N}$. Then*

$$m \leq n \iff \exists b \in \mathbb{N}\,(n = m + b).$$

Proof. First suppose $m \leq n$. We wish to find a b such that $n = m + b$. By Definition 2, we have two cases: (i) $m < n$ and (ii) $m = n$. In case (i) the existence of b follows from Definition 5. In case (ii) we can take $b = 0$. In either case, we have a suitable $b \in \mathbb{N}$.

Next suppose $n = m + b$ for some b. We must show that $m \leq n$. If $b = 0$ then $m = n$, which implies $m \leq n$ by Definition 2. If $b \neq 0$ then $m < n$ by Definition 5. Thus $m \leq n$ by Definition 2. □

3. Basic properties of the order on $\mathbb{N}$

In Chapter 1 we defined the set $\mathbb{N}^+$ of positive natural numbers in terms of the condition $n \neq 0$. However, most people use the condition $n > 0$. Both work for $\mathbb{N}$, but in fact $n > 0$ is the right condition for other number systems. We didn't use the condition $n > 0$ in Chapter 1 because the relation $>$ had not been defined yet. We now show that both conditions are equivalent for $\mathbb{N}$, so it doesn't matter which you use to define positive natural numbers.

Theorem 12. *Suppose $n \in \mathbb{N}$. Then $n \neq 0$ if and only if $n > 0$.*

Proof. Part 1. Suppose that $n \neq 0$. Observe that $n = 0 + b$ where $b = n$. So $b \neq 0$. Thus, by Definition 5, $n > 0$.

Part 2. Suppose that $n > 0$. By Definiton 5, we have $n = 0 + b$ where $b \neq 0$. Thus $n = b$. Since $b \neq 0$, we have $n \neq 0$. □

Corollary 13. *A natural number n is positive if and only if $n > 0$, and*

$$\mathbb{N}^+ = \{n \in \mathbb{N} \mid n \neq 0\} = \{n \in \mathbb{N} \mid n > 0\}.$$

The following, like many results in this chapter, is so ingrained into our thinking that it is easy to forget to prove it:

Theorem 14. *Let $n \in \mathbb{N}$. There is no $x \in \mathbb{N}$ such that $n < x < n + 1$. In other words, are no natural numbers between n and $n + 1$.*

Proof. (By contradiction). Suppose that $n < x$ and $x < n+1$. Since $n < x$, there is a positive b such that $x = n+b$. Since $x < n+1$, there is a positive c such that $n+1 = x+c$. Since $b \neq 0$, it has a predecessor. So $b = \sigma d = 1+d$ for some $d \in \mathbb{N}$. Putting this together gives

$$x = n + b = n + (1 + d) = (n + 1) + d = (x + c) + d = x + (c + d).$$

Thus, $0 + x = (c + d) + x$. By the cancellation law, $c + d = 0$. By a result in Chapter 1, we must have $c = d = 0$. But $c = 0$ is a contradiction. □

Exercise 8. Identify the results and definitions used in the above proof.

Exercise 9. Show that the only $x \in \mathbb{N}$ such that $x < 2$ are $x = 0$ and $x = 1$. Hint: use addition facts to show $x = 0$ and $x = 1$ work. Now suppose that $x \neq 0, 1$. Divide into cases: $x < 1$ and $1 < x$.

Theorem 15. *Suppose that $x, y, z \in \mathbb{N}$. If $x \leq y$ then $xz \leq yz$.*

Proof. Suppose that $x \leq y$. By Theorem 11 there is a $b \in \mathbb{N}$ where $y = x + b$. By the distributive law, $yz = (x + b)z = xz + bz$. Thus $xz \leq yz$ by Theorem 11. □

Exercise 10. Prove the following theorem using Definition 5 and Theorem 11.

Theorem 16. *Suppose that $x, y, z \in \mathbb{N}$.*
Then $x < y$ if and only if $x + z < y + z$.
Similarly, $x \leq y$ if and only if $x + z \leq y + z$.

Exercise 11. The remaining results of this section are given with sketchy proofs. Rewrite them in a more complete, organized form.

Theorem 17. *Suppose that $x, y, z \in \mathbb{N}$ where $z > 0$. If $x < y$ then $xz < yz$.*

Proof. Assume $x < y$. There is a $b \in \mathbb{N}$ such that $z = b + 1$. So

$$xz = x(b + 1) = xb + x \leq yb + x < yb + y = y(b + 1) = yz.$$

□

Theorem 18. *Suppose that $x, y, z \in \mathbb{N}$. If $xz < yz$ then $x < y$.*

Proof. Suppose $xz < yz$, but that $x < y$ fails. Then $y \leq x$ by the trichotomy law. Now use Theorem 15 to derive a contradiction. □

Exercise 12. Suppose $m_1, m_2, n_1, n_2 \in \mathbb{N}$, $m_1 < m_2$ and $n_1 < n_2$. Show $m_1 + n_1 < m_2 + n_2$ and $m_1 n_1 < m_2 n_2$. Hint: for the last inequality, it helps to first show that $m_2 > 0$.

4. Cancellation law for multiplication

In Chapter 1 we proved the cancellation law for addition, but we postponed the multiplicative cancellation law until we developed properties of $<$. These properties allow for a quick and easy proof of the law.

Theorem 19 (Cancellation Law for Multiplication)**.** *Suppose $x, y, z \in \mathbb{N}$. If $xz = yz$ and $z \neq 0$ then $x = y$.*

Proof. By the trichotomy law, either $x = y$, $x < y$ or $y < x$. The last two cases lead to contradictions via Theorem 17 and the trichotomy law. Thus $x = y$. □

Another important theorem is the following:

Theorem 20. *Suppose $m, n \in \mathbb{N}$. If $mn = 0$ then $m = 0$ or $n = 0$.*

Proof. Assume $mn = 0$, but that m and n are both nonzero. Since $mn = 0$ and $0n = 0$, we have $mn = 0n$. Thus $m = 0$ by the cancellation law for multiplication. This is a contradiction. □

Exercise 13. Prove that $\mathbb{N}^+$ is closed under multiplication. Show this as a corollary to the above theorem.

Exercise 14. Suppose $n, B \in \mathbb{N}$ where $B \neq 0$. Show $B^n \neq 0$ for all $n \in \mathbb{N}$.

5. The Maximum principle and the well-ordering property

Above we established that $\mathbb{N}$ is linearly ordered by $<$. Several other number systems, including $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$, are also linearly ordered. We now consider a general definition of a linearly ordered set in order to describe the concepts of *minimum* and *maximum*:

Definition 6 (Linearly Ordered Set)**.** An *linearly ordered set* is a set U with a designated strict linear order relation $<$.

Remark 5. If we designate the order relation $<$ from Definition 5, then $\mathbb{N}$ becomes a linearly ordered set. In later chapters we will see how $\mathbb{Z}$, $\mathbb{Q}$ and $\mathbb{R}$ can be considered to be linearly ordered sets. (There are linear orders on $\mathbb{C}$ but they are all somewhat artificial, so it is not useful for us to think of $\mathbb{C}$ as a linearly ordered set: there is no order relation that we would want to specially designate for $\mathbb{C}$.)

Remark 6. From $<$ we define $\leq, >, \geq$ as in Section 1. So, in a sense, a linearly ordered set has four designated order relations $(<, \leq, >, \geq)$.

Remark 7. If S is a subset of an ordered set U, then S is itself an ordered set. We just restrict the designated strict linear order to the subset S. Thus any subset of $\mathbb{N}$ is an ordered set.

Definition 7. Let S be a subset of an ordered set U. An element $b \in U$ is called a *lower bound* of S if $b \leq x$ holds for all $x \in S$. An element $B \in U$ is called an *upper bound* of S if $x \leq B$ holds for all $x \in S$.

Definition 8. Let S be a subset of an ordered set U. An element $m \in S$ is called a *minimum* of S if $m \leq x$ holds for all $x \in S$. An element $M \in S$ is called a *maximum* of S if $x \leq M$ holds for all $x \in S$.

Exercise 15. How do the above two definitions differ? Use the answer to this question to prove the following:

Lemma 21. *Suppose that S is a subset of an ordered set U. If b is a lower bound of S that is also an element of S then b is a minimum of S. Similarly, if B is an upper bound of S that is also an element of S then B is a maximum of S.*

Warning. Not all subsets of ordered sets have a minimum and a maximum. For example, if $U = \mathbb{N}$ then $S = U$ has no maximum. In a later chapter we will describe intervals such as $S = (0, 1]$ in $\mathbb{Q}$ that have no minimum. The

number 0 is not a minimum of $(0, 1]$ since it is not in S, but 0 is a lower bound for S.

Existence may fail, but if existence holds then uniqueness must as well:

Theorem 22. *The minimum of U, if it exists, is unique. The maximum of U, if it exists, is unique.*

Exercise 16. Prove the above theorem.

Warning. In contrast, upper and lower bounds are not necessarily unique.

Exercise 17. What is the minimum of $\mathbb{N}$? Show that $\mathbb{N}$ has no maximum.

Although $\mathbb{N}$ itself has no maximum, we will now prove that every nonempty subset of $\mathbb{N}$ with an upper bound does have a maximum. This is in contrast to other number systems we will see such as $\mathbb{Q}$ and $\mathbb{R}$ where bounded subsets do not always have maxima.

Informal Exercise 18. Find a nonempty subset of $\mathbb{Q}$ that has an upper bound, but no maximum.

Theorem 23 (Maximum Principle)**.** *Suppose W is a nonempty subset of $\mathbb{N}$ with an upper bound. Then W has a maximum.*

Proof. Consider the following set

$$T \stackrel{\text{def}}{=} \{B \in \mathbb{N} \mid \text{every nonempty } S \subseteq \mathbb{N} \text{ with upper bound } B \text{ has a maximum}\}$$

So $B \in T$ means that *any* subset of $\mathbb{N}$ bounded by B has a maximum. Our first goal is to use induction to show that all natural numbers are in T. Once we have done this, proving the theorem will be easy.

We begin by showing the base case: $0 \in T$. To do so, let S be any nonempty subset of $\mathbb{N}$ with upper bound 0. Since S is nonempty, it has an element k. Since 0 is an upper bound for S, $k \leq 0$. Since k is a natural number $0 \leq k$. These two inequalities mean that $k = 0$. So the upper bound 0 is an element of S. By Lemma 21, this means that 0 is the maximum of S, and so S possesses a maximum. This is true of all subsets bounded by 0, so $0 \in T$.

Now suppose $n \in T$. We wish to show that $n + 1 \in T$. In other words, we consider an arbitrary subset $S \subseteq \mathbb{N}$ with upper bound $n + 1$, and we must show that S has a maximum. We divide into two cases: $n + 1 \in S$ and $n + 1 \notin S$.

If $n + 1 \in S$ we are done by Lemma 21.

If $n + 1 \notin S$ then we claim that n must also be an upper bound of S. If not, then there is an element $k \in S$ such that $n < k$. Since $n+1$ is an upper bound, but not an element of S, we have $k < n + 1$. So

$$n < k < n + 1$$

which contradicts Theorem 14 that there are no elements in $\mathbb{N}$ between n and $n + 1$. So our claim that n is an upper bound of S is justified. Since $n \in T$, we conclude that S has a maximum.

In both cases ($n+1 \in S$ and $n+1 \notin S$) we have shown that S has a maximum. By definition of T, this yields $n+1 \in T$ as desired.

Now we can invoke the induction axiom to conclude that $T = \mathbb{N}$. We now use this fact to prove the main result. Let W be any nonempty subset of $\mathbb{N}$ with an upper bound. Let B be an upper bound. Note that $B \in T$ since $T = \mathbb{N}$. This means that every subset with upper bound B possesses a maximum. In particular, W itself has a maximum. □

We now come to a key concept of this chapter:

Definition 9. An ordered set U is said to be *well-ordered* if every non-empty subset $S \subseteq U$ has a minimum.

Warning. Showing that $\mathbb{N}$ has a minimum is not enough to prove it is well-ordered. You must show that every nonempty subset of $\mathbb{N}$ has a minimum. Of course, different subsets can have different minima.

Theorem 24. *The set of natural numbers $\mathbb{N}$ is well-ordered.*

Proof. We will not prove this by induction, but instead will prove it as a corollary of the Maximum Principle. Let S be a nonempty subset of $\mathbb{N}$ that we want to show has a minimum. Let L be the set of lower bounds of S.

Observe that $0 \in L$ (why?), so L is nonempty. Since S is nonempty, there is an element $k \in S$. Observe that k is an upper bound of L. By the Maximum Principle, the set L has a maximum element n. In other words, there is a maximum lower bound of S. In particular, n is a lower bound of S, and we would like to show that n is a minimum of S as well. This requires showing that $n \in S$.

In order to show $n \in S$, suppose otherwise. Since n is a lower bound, this means $n < s$ for all $s \in S$. Since there are no elements strictly between n and $n+1$, this means that $n+1 \leq s$ for all $s \in S$. Observe that this means $n+1 \in L$. This contradicts the definition of n as the maximum of L. Thus $n \in S$, and so n is the minimum of S by Lemma 21. □

Exercise 19. Justify the observations in the above proof: (i) $0 \in L$, (ii) k is an upper bound of L, and (iii) if $n \notin S$ then $n+1 \in L$.

Informal Exercise 20. Is the set of nonnegative rational numbers well-ordered? Is the set of integers $\mathbb{Z}$ well-ordered?

6. Subtraction

In Chapter 1 we considered addition and multiplication for the natural numbers, but we did not consider subtraction. This is because, $n-m$ is not defined for all $m, n \in \mathbb{N}$. However, now we have an order relation on $\mathbb{N}$, and when $n \geq m$ we *can* define subtraction. (In Chapter 4 we will introduce negative integers, and then we will be able to define $n-m$ for all integers n and m.)

Recall from Theorem 11 that $n \geq m$ if and only if there is a $b \in \mathbb{N}$ such that $n = m + b$. It turns out that this b is unique:

Lemma 25. *Let $n, m \in \mathbb{N}$. If $n \geq m$ then there is a unique $b \in \mathbb{N}$ such that $n = m + b$.*

Exercise 21. Prove the above lemma.

In mathematics, when we defined a new term such as $n - m$, we need to identify a specific object that the term will reference. Both existence and uniqueness are important in identifying a specific object with a given property. For example we can specify b with the property $n = m+b$ when $n \geq m$ since such a b exists, and there is no ambiguity since b is unique.

Definition 10. Let $m, n \in \mathbb{N}$ be such that $n \geq m$. Then $n - m$ is defined to be the $b \in \mathbb{N}$ such that $n = m + b$. We call $n - m$ the *difference* of n and m, and call $-$ the *subtraction operation*.

The subtraction operation is not defined for all $(n, m) \in \mathbb{N} \times \mathbb{N}$, but only for the subset consisting of pairs where $n \geq m$. This means that subtraction is *not* a binary operation on $\mathbb{N}$. (In Chapter 3 we will define $n - m$ without assuming $n \geq m$, but the result will not always be in $\mathbb{N}$. We will see that subtraction *is* a binary operation on $\mathbb{Z}$. Later we will see that it is a binary operation on $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$ as well.)

Directly from the definition we have the following.

Theorem 26 (Basic law of subtraction)**.** *Suppose $m, n, b \in \mathbb{N}$ and $n \geq m$. Then $n = m + b$ if and only if $b = n - m$.*

Theorem 27. *Let $n \in \mathbb{N}$. Then $n - n = 0$.*

Proof. Since $n = n + 0$, the conclusion $n - n = 0$ follows from Theorem 26. □

Exercise 22. Prove the following four theorems as consequences of Theorem 26.

Theorem 28. *Given $n, m \in \mathbb{N}$ with $n \geq m$, if $n - m = 0$ then $n = m$.*

Theorem 29. *Given $n \in \mathbb{N}$, then $n - 0 = n$.*

Theorem 30. *Suppose $m, n \in \mathbb{N}$ with $n \geq m$. Then $m + (n - m) = n$.*

Theorem 31. *Suppose $y \leq x$ and $z \leq x$ where $x, y, z \in \mathbb{N}$. Then $x - y = z$ if and only if $x - z = y$.*

Remark 8. Parentheses are often required to determine the meaning of an expression involving subtraction. For example, $(9 - 8) - 2$ is not defined since $1 < 2$. However, $9 - (8 - 2)$ is defined, and is equal to 3. (Note that $8 - 2 = 6$ since $8 = 2 + 6$, and $9 - 6 = 3$ since $9 = 6 + 3$.)

When parentheses are not explicitly written, we follow the usual rules for grouping. One rule we will adopt is that when we are given terms linked by $+$ and $-$, we perform our operations left to right. For example

$$a + b + c - d + e - f + (g + h) - (i - j) + k.$$

is really

$$\Big(\Big(\big(\big(\big(\big((a+b)+c\big)-d\big)+e\big)-f\big)+(g+h)\Big)-(i-j)\Big)+k.$$

Even though parentheses are usually necessary when using subtraction, the following shows a situation where parentheses can be moved.

Theorem 32. *Suppose $x, y, z \in \mathbb{N}$ are such that $z \leq y$. Then*

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

Proof. Let $c = x+(y-z)$. Observe that

$$c+z = \big(x+(y-z)\big)+z = x+\big((y-z)+z\big) = x+y$$

by the associative and commutative laws of addition (Chapter 1) and Theorem 30. Thus $z+c = x+y$ by the commutative law. By the basic law of subtraction $c = (x+y)-z$. The result follows. □

Here is an application of the above law:

Theorem 33. *Suppose $m, n, c \in \mathbb{N}$. If $n \geq m$ then $n+c \geq m+c$ and*

$$n-m = (n+c)-(m+c).$$

Proof. The first part follows from Theorem 16. For the second part:

$$\begin{aligned}
(m+c)+(n-m) &= ((m+c)+n)-m && \text{(Theorem 32)}\\
&= (n+(c+m))-m && \text{(Comm law: Ch.1, twice)}\\
&= ((n+c)+m)-m && \text{(Assoc law: Ch.1)}\\
&= (n+c)+(m-m) && \text{(Theorem 32)}\\
&= (n+c)+0 && \text{(Theorem 27)}\\
&= n+c && \text{(Rule from Ch.1)}
\end{aligned}$$

The result now follows from the basic law of subtraction. □

7. The set $\{1, \ldots, n\}$

The purpose of this section is to define and develop basic properties of the set $\{1, \ldots, n\}$ where $n \in \mathbb{N}$. This set will be important in the next chapter when we use it to define the cardinality of a finite set.

Definition 11. Let $n \in \mathbb{N}$ Then $\{1, \ldots, n\}$ is defined as the set

$$\{x \in \mathbb{N} \mid 1 \leq x \leq n\}.$$

More generally if $m, n \in \mathbb{N}$ the $\{m, \ldots, n\}$ is defined as $\{x \in \mathbb{N} \mid m \leq x \leq n\}$. Warning: this is the empty set if $m > n$.

We allow common notational variants. For example, $\{1, 2, \ldots, n\}$ is also defined as $\{x \in \mathbb{N} \mid 1 \leq x \leq n\}$.[1]

[1]Warning: the notation $\{1, 2, \ldots, n\}$ might suggest to the reader that $n > 2$. To be safe, we should always mention any assumptions about the size of n.

Here are facts from set theory, repeated for convenience:

$$\{a\} = \{x \mid x = a\}$$
$$\{a, b\} = \{x \mid (x = a) \vee (x = b)\}.$$
$$\{a, b, c\} = \{x \mid (x = a) \vee (x = b) \vee (x = c)\}.$$

Theorem 34. *Let $m \in \mathbb{N}$. The set $\{m, \ldots, m\}$ is equal to $\{m\}$.*

Proof. Suppose $x \in \{m, \ldots, m\}$. By Definition 11, $m \leq x \leq m$. However, $m \leq x$ and $x \leq m$ imply $x = m$. From set theory we know that $x = m$ implies $x \in \{m\}$. We conclude that $\{m, \ldots, m\} \subseteq \{m\}$.

Suppose $x \in \{m\}$. By basic set theory, this means $x = m$. Since $x = m$ we have both $m \leq x$ and $x \leq m$. In other words, $m \leq x \leq m$. So, by Definition 11, $x \in \{m, \ldots, m\}$. Thus $\{m\} \subseteq \{m, \ldots, m\}$. □

Theorem 35. *Let $n \in \mathbb{N}$. The set $\{n, \ldots, n+1\}$ is equal to $\{n, n+1\}$.*

Exercise 23. Prove the above theorem.

The next theorem one might want to show is that

$$\{n, \ldots, n+2\} = \{n, n+1, n+2\}.$$

After this, one might want a theorem concerning $\{n, \ldots, n+3\}$, and so on. The next theorem shows a key relationship that helps make it easier to prove such results.

Theorem 36. *If $m, n \in \mathbb{N}$ where $m \leq n$, then*

$$\{m, \ldots, n+1\} = \{m, \ldots, n\} \cup \{n+1\}.$$

Furthermore, $n + 1 \notin \{m, \ldots, n\}$ so the union is disjoint.

Proof. First we prove $\{m, \ldots, n+1\} \subseteq \{m, \ldots, n\} \cup \{n+1\}$. So suppose that $x \in \{m, \ldots, n+1\}$. Then, by definition, $m \leq x$ and $x \leq n+1$. So either $x = n+1$ or $x < n+1$.

Case 1: $x = n+1$. So by basic set theory, $x \in \{m, \ldots, n\} \cup \{n+1\}$.

Case 2: $x < n+1$. Observe that $n < x$ implies $n < x < n+1$ which cannot happen (Theorem 14). Thus $x \leq n$. We know that $m \leq x$, so $m \leq x \leq n$. Thus $x \in \{m, \ldots, n\}$. Since $\{m, \ldots, n\} \subseteq \{m, \ldots, n\} \cup \{n+1\}$, we have $x \in \{m, \ldots, n\} \cup \{n+1\}$ as well.

So in either case, $x \in \{m, \ldots, n\} \cup \{n+1\}$. Thus

$$\{m, \ldots, n+1\} \subseteq \{m, \ldots, n\} \cup \{n+1\}.$$

Next we will prove that $\{m, \ldots, n\} \cup \{n+1\} \subseteq \{m, \ldots, n+1\}$. So suppose that $x \in \{m, \ldots, n\} \cup \{n+1\}$ By definition of union, we have two cases.

Case 1: $x \in \{m, \ldots, n\}$. In this case $m \leq x \leq n$. But $n < n+1$ by Theorem 5, Thus $x < n+1$ by (mixed) transitivity. In particular $x \leq n+1$. So $m \leq x \leq n+1$, thus $x \in \{m, \ldots, n+1\}$.

Case 2: $x \in \{n+1\}$ In other words, $x = n+1$. Now $n < n+1$, by Theorem 5. In particular, $n \leq x$. By hypothesis, $m \leq n$. Thus $m \leq x$ by

transitivity. Trivially $x \leq n+1$, since $x = n+1$. So $m \leq x \leq n+1$. We conclude that $x \in \{m, \ldots, n+1\}$.

In either case, $x \in \{m, \ldots, n+1\}$. Thus

$$\{m, \ldots, n\} \cup \{n+1\} \subseteq \{m, \ldots, n+1\}.$$

Combining this with the previous inclusion, we conclude the sets are equal.

Finally, we show $n+1 \notin \{m, \ldots, n\}$. Suppose otherwise: $m \leq n+1 \leq n$. Then $n+1 \leq n$. However, by Theorem 5, $n < n+1$. This contradicts trichotomy. □

Exercise 24. Use the above to give another proof that the set $\{n, \ldots, n+1\}$ is just $\{n, n+1\}$. Hint: $\{a\} \cup \{b\} = \{a, b\}$ by basic set theory.

Exercise 25. Show $\{1, \ldots, 3\} = \{1, 2, 3\}$. Hint: $\{a\} \cup \{b\} \cup \{c\} = \{a, b, c\}$ by basic set theory.

8. Simple recursion

It is common to define a function $g : \mathbb{N} \to S$ by recursive equations. These are equations that define $g(n)$ in terms of other values $g(m)$ of the same function g. This seems circular, but it is not since we will always require $m < n$.

For example, suppose we want to define a function $g : \mathbb{N} \to \mathbb{N}$ by the equations

$$g(0) = 1, \qquad \text{and} \qquad g(n+1) = 2g(n) + 1.$$

These equations force $g(0) = 1, \quad g(1) = 2g(0) + 1 = 3,$

$$g(2) = 2g(1) + 1 = 2 \cdot 3 + 1 = 7, \quad \text{and so on.}$$

It seem clear that these equations define a unique function $g : \mathbb{N} \to \mathbb{N}$ given our intuitive idea of the natural numbers. How do we prove this? Uniqueness is not hard to show, but what about existence?

Exercise 26. Show uniqueness of g using induction.

The iteration theorem from Chapter 1 gives existence quite easily. First observe that the equation $g(n+1) = 2g(n) + 1$ makes the next value a function of the previous value. The function f that gives the next value is given by the rule $x \mapsto 2x+1$. You get the values of g by iterating f starting with 1. So define g by the equation $g(n) = f^n(1)$.

Exercise 27. Let $g\colon \mathbb{N} \to \mathbb{N}$ be defined by the rule $g(n) = f^n(1)$ where f is as above. Show that $g(0) = 1$ based on the fact that f^0 is the identity function. Show that $g(n+1) = 2g(n)+1$ based on the fact that $f^{n+1} = f \circ f^n$.

The above discussion generalizes to the following theorem.

Theorem 37 (Simple Recursion)**.** *Let S be a set. Suppose that $f : S \to S$ and $a \in S$ are given. Then there is a unique function $g : \mathbb{N} \to S$ satisfying the equations*

$$g(0) = a, \qquad \textit{and} \qquad g(n+1) = f(g(n)).$$

Furthermore, g is given by $g(n) = f^n(a)$.

Exercise 28. Prove the above theorem.

9. More advanced recursion

A famous function defined by recursion is the *Fibonnaci function*. This is defined by the recursive equations:

$$F(0) = 0, \quad F(1) = 1, \quad F(n+2) = F(n) + F(n+1).$$

The difference between this and simple recursion is that, in general, a value of F depends not only on the previous value of F, but the previous *two* values of F. Note that the equations force

$$F(0) = 0, \quad F(1) = 1, \quad F(2) = 1 + 0 = 1, \quad F(3) = 1 + 1 = 2,$$

$$F(4) = 1 + 2 = 3, \quad F(5) = 2 + 3 = 5, \quad F(6) = 3 + 5 = 8,$$

and so on. It is obvious that these equations define a unique function $\mathbb{N} \to \mathbb{N}$ given our intuitive idea of the natural numbers. However, can we prove existence and uniqueness given what we have rigorously established?

To use iteration to prove the existence of the Fibonnaci function we use a trick: we switch to $\mathbb{N} \times \mathbb{N}$. We are interested in pairs (x, y) where x is a given Fibonnaci number, and y is the next Fibonnaci number. We also consider the function $\theta : \mathbb{N} \times \mathbb{N} \to \mathbb{N} \times \mathbb{N}$ defined by the rule $(x, y) \mapsto (y, x + y)$ which advances us from one pair of adjacent Fibonnaci numbers to the next pair of Fibonnaci numbers. For example, $\theta(3, 5) = (5, 8)$.

The initial pair is $(0, 1)$, so consider $\theta^n(0, 1)$ as the nth pair. Define $F(n)$ to be the first coordinate of $\theta^n(0, 1)$.

Exercise 29. Prove that $F(n+1)$ is the second coordinate of $\theta^n(0, 1)$. Hint: by definition of $F(n)$ we have $\theta^n(0, 1) = \big(F(n), y\big)$ for some $y \in \mathbb{N}$. Now apply θ to both sides of this equation to conclude that $y = F(n + 1)$.

From the above exercise, we have that $\theta^n(0, 1) = \big(F(n), F(n + 1)\big)$ for all $n \in \mathbb{N}$. The special case $n = 0$ gives us $\theta^0(0, 1) = \big(F(0), F(1)\big)$.

Exercise 30. Show $F(0) = 0$ and $F(1) = 1$.

Exercise 31. Show that $F(n + 2) = F(n) + F(n + 1)$ for all $n \in \mathbb{N}$. Hint: apply θ to both sides of the equation $\theta^n(0, 1) = \big(F(n), F(n+1)\big)$. Now look at the second coordinate of both sides.

We now have the existence of F. What about the uniqueness?

Exercise 32. Show that there is a unique solution $F : \mathbb{N} \to \mathbb{N}$ to the equations

$$F(0) = 0, \quad F(1) = 1, \quad F(n + 2) = F(n) + F(n + 1).$$

Hint: suppose F_1 and F_2 are two distinct solutions. Let S be the set of n such that $F_1(n) \neq F_2(n)$. So S has a least element m by the well-ordering theorem. Observe that $m \neq 0$ and $m \neq 1$. Conclude that $m \geq 2$. Now derive a contradiction.

We end with a different sort of recursion. The following equations defines the so-called *triangular numbers*:

$$T(0) = 0, \quad T(n+1) = (n+1) + T(n).$$

The difference between this and simple recursion is that $T(n+1)$ is not a function of $T(n)$ alone, but also depends on n. In other words, you need to know both $T(n)$ and n (or $n+1$) in order to find $T(n+1)$. Note that the equations force

$$T(0) = 0, \quad T(1) = 1 + 0 = 1, \quad T(2) = 2 + 1 = 3, \quad T(3) = 3 + 6 = 6,$$

and so on.[2] It is obvious, from our intuitive idea of the natural numbers, that these equations define a unique function $T : \mathbb{N} \to \mathbb{N}$. How do we prove this?

To use iteration to prove the existence of the function $T : \mathbb{N} \to \mathbb{N}$ we use a trick: we work in $\mathbb{N} \times \mathbb{N}$. We are interested in pairs (x, y) such as $(3, 6)$ where y is the xth triangular number. We also consider the function

$$\psi : \mathbb{N} \times \mathbb{N} \to \mathbb{N} \times \mathbb{N}$$

defined by the rule $(x, y) \mapsto (x+1, (x+1)+y)$ which advances us from one triangular number to the next. For example, $\psi(2, 3) = (3, 6)$.

We then define $T(n)$ to be the second coordinate of $\psi^n(0, 0)$, and prove it satisfies the desired equations.

Exercise 33. Prove, by induction, that the first coordinate of $\psi^n(0, 0)$ is n. Thus

$$\psi^n(0, 0) = \big(n, T(n)\big).$$

Exercise 34. Use the equation $\psi^n(0, 0) = \big(n, T(n)\big)$ to show $T(0) = 0$.

Exercise 35. Show $T(n+1) = (n+1) + T(n)$. Hint: apply ψ to both sides of the equation $\psi^n(0, 0) = \big(n, T(n)\big)$.

We see from these exercises the existence of a solution $T : \mathbb{N} \to \mathbb{N}$ to the equations

$$T(0) = 0, \quad T(n+1) = (n+1) + T(n).$$

Here is a generalization of the triangular number example:

Theorem 38. *Let S be a set, c an element of S, and $g : \mathbb{N} \times S \to S$ a function. Then there is a unique function $f : \mathbb{N} \to S$ satisfying the equations*

$$f(0) = c, \quad f(n+1) = g\big(n, f(n)\big).$$

for all $n \in \mathbb{N}$.

Proof. Let $\gamma : \mathbb{N} \times S \to \mathbb{N} \times S$ be defined by the rule $(n, x) \mapsto \big(n+1, g(n, x)\big)$. Define $f(n)$ to the the second coordinate of $\gamma^n(0, c)$. This function can be shown to satisfy the equations (see above discussion). Induction can be used to show uniqueness. □

[2] It turns out that $T(n) = n(n+1)/2$, so once we have developed division, we do not need to define T recursively. However, the recursive definition captures the idea of a triangle better than the formula $T(n) = n(n+1)/2$.

Definition 12. The *factorial function* $f\colon \mathbb{N} \to \mathbb{N}$ is defined by using the above theorem with $S = \mathbb{N}$, $c = 1$, and $g(n, m) = (n+1)\cdot m$. In other words, it is the solution to the recursive equations

$$f(0) = 1, \quad f(n+1) = (n+1)f(n).$$

We write $n!$ for $f(n)$. So

$$0! = 1, \quad (n+1)! = (n+1)n!.$$