1.1 The Natural Numbers

The elements of the set of natural numbers:

$$\mathbb{N} = \{1, 2, 3, 4, 5, ...\}$$

are the numbers we use for counting. They come equipped with an **ordering**:

$$1 < 2 < 3 < 4 < ...$$

and they also come equipped with the:

Well-ordered axiom: Every set of natural numbers except the empty set has a smallest element.

Note: This is an axiom, meaning we will accept it without demanding a proof. (Why do you suppose mathematicians are willing to accept this?)

From the well-ordered axiom we may deduce the:

Principle of induction: If S is a subset of $\mathbb{N}$, such that:

(i) $1 \in S$ and

(ii) whenever $n \in S$, the next number after n is also an element of S

then S is equal to $\mathbb{N}$, the set of all natural numbers.

Note: This is **not** given as an axiom, so we have to prove it!

Proof: Consider the complementary set S^c whose elements are the natural numbers that are **not** elements of S. This is also a set of natural numbers, to which we will apply the well-ordered axiom. In other words, either S^c has a smallest element or else it is the empty set. Let's suppose S^c has a smallest element. That element can't be 1 (because $1 \in S$ by (i)) and it can't be anything bigger than 1 (because whenever S^c contains a number bigger than 1, then S^c also contains the number immediately **before** it by (ii)). So **no** natural number can be the smallest element of S^c, so S^c is empty, which is the same as saying that S is the set of all natural numbers.

We can do a lot with the principle of induction. Any time you might be tempted to write the abbreviation "etc." or "..." in a proof or definition, then you probably should use the principle of induction. For example, we define the addition of 1 to any natural number m as follows:

Definition of Addition of 1: $m + 1$ is the next number after m.

But we'd also like to be able to define addition of anything as follows:

$$m + 2 \text{ is the next number after } m + 1,$$

$$m + 3 \text{ is the next number after } m + 2,$$

$$\text{etc.}$$

This is a job for the principle of induction.

Strategy for making definitions by induction.

A function f whose domain is $\mathbb{N}$ may be defined in two steps:

(i) Define $f(1)$.

(ii) Define each $f(n+1)$, possibly using a previous definition of $f(n)$.

The domain of the function f is then a subset of $\mathbb{N}$ that contains 1 by (i), and whenever n is in the domain, then $n+1$ is also in the domain by (ii). Therefore the domain of f is $\mathbb{N}$, by the principle of induction!

Definition of Addition: $m+n$ is defined for all m and n as follows:

(i) $m+1$ is, as above, defined to be the next number after m.

(ii) Each $m+(n+1)$ is defined to be the next number after $m+n$.

Think of m as a fixed natural number and n as a variable. Then (i) and (ii) define the function $f(n) = m+n$ for all values of n by the strategy above for definitions by induction.

Clarification: Since we are regarding m as fixed while n is a variable, the two letters play quite different roles. Because of this, we need to prove annoying things like the commutativity of addition(!) If we had thought of them both as variables, we would not have been able to use the principle of induction to define addition because $f(m,n) = m+n$ would have been a function of two variables!

Next we turn to proofs by induction. A **mathematical sentence** P is an (ordinary) sentence that is definitely either true or false. For example:

- "There are 5 days in a week" is a false mathematical sentence,
- "$14 > 13$" is a true mathematical sentence, and
- "$5+2=8$" is another false mathematical sentence, but
- "It was a dark and stormy night" is too vague to be mathematical.

$P(n)$ will stand for an ordinary sentence that may contain the variable n (but no other variables) and which becomes a mathematical sentence whenever n is given a (natural number) value. These are much like functions $f(n)$, which contain the variable n and become a number whenever n is given a natural number value. Sentences $P(n)$ may be true for all values of n, or they may be only true for some (or no) values. For example:

- "There are n days in a week" is only true for the value $n=7$. But
- "$n+1 > n$" is true for all values of n, and
- "$n+1=n$" is never true (but it is still of the form $P(n)$), while
- "$n=m$" is only of the form $P(n)$ when a value for m is chosen.

Suppose we are given $P(n)$, which we think should be true for all n.

Strategy for proving $P(n)$ is true for all n by induction.

This may be done in two steps:

(i) Prove that the sentence $P(1)$ is true.

(ii) Prove that each sentence $P(n+1)$ is true, possibly making use of the assumed truth of $P(n)$.

Define S to the set of natural numbers that make $P(n)$ true. $1 \in S$ by (i), and whenever $n \in S$, then $n+1 \in S$, by (ii). Thus $S = \mathbb{N}$ by the principle of induction, so proving (i) and (ii) proves that the sentences $P(n)$ are all true.

Associativity Law of Addition:

$$(l+m)+n = l+(m+n)$$

for all natural numbers l, m, n.

Proof: Think of l and m as fixed. We follow the strategy for a proof by induction to prove, for all n, the associativity sentences:

$$\text{“}(l+m)+n = l+(m+n)\text{”}$$

which we'll call $P(n)$.

(i) By addition definition (i), $(l+m)+1$ is the next number after $l+m$, and by addition definition (ii), $l+(m+1)$ is also the next number after $l+m$. Since there is only one next number after $l+m$, we get:

$$(l+m)+1 = l+(m+1)$$

That is, $P(1)$ is proved to be true.

More Clarification: In the proof of (i) above, $l+m$ plays the role of m in addition definition (i), and l and m play the roles of m and n, respectively, in addition definition (ii). This scrambling of letters is unfortunate in these proofs, but it is also pretty much unavoidable. When you are checking these proofs and constructing your own, you might want to keep careful track of such scrambles. One way of keeping track would be with footnotes. As an example, I've footnoted the scrambles for you in the remainder of this proof.

(ii) We need to prove each $P(n+1)$:

$$(l+m)+(n+1) = l+(m+(n+1))$$

allowing ourself to assume $P(n)$:

$$(l+m)+n = l+(m+n)$$

We will do this by taking the next numbers after each side of equation $P(n)$ (which must then be equal to each other!). Applying addition definition (ii)[1] to the left side, we see that:

$$(l+m)+(n+1) \text{ is the next number after } (l+m)+n$$

On the other hand, applying addition definition (ii)[2] to the right side of the equation, we see that $l+((m+n)+1)$ is the next number after $l+(m+n)$. Moreover, using $P(1)$ above[3], which we have already proved(!) and substitution, we get $l+((m+n)+1) = l+(m+(n+1))$, so:

$$l+(m+(n+1)) \text{ is the next number after } l+(m+n)$$

and putting these together, we do indeed get $(l+m)+(n+1) = l+(m+(n+1))$ completing part (ii) of the proof.

Thus our proof by induction strategy proves that the associativity sentence is true for every n. And since l and m could be anything, we have proved the associative law of addition!

You will prove the commutative law of addition (in the exercises).

Definition of Multiplication: $m \times n$ is defined for all m and n as follows:

(i) $m \times 1 = m$.

(ii) Each $m \times (n+1) = m \times n + m$.

Think of m as a fixed natural number and n as a variable. Then $f(n) = m \times n$ (also written mn) is defined for all n by induction.

Note: By writing $m \times n + m$ or $mn + m$, we mean, of course, $(m \times n) + m$, following the standard rules for the order of arithmetic operations.

Multiplication distributes with addition:

$$(l+m) \times n = l \times n + m \times n$$

for all natural numbers l, m, n.

Proof: Think of l and m as fixed, and apply the induction strategy to prove:

$$\text{“}(l+m) \times n = l \times n + m \times n\text{”}$$

(the distributivity sentence, which we'll call $P(n)$) for all n.

(i) Two applications of multiplication definition (i) give $P(1)$:

$$(l+m) \times 1 = l+m = l \times 1 + m \times 1$$

[1] Here, $l+m$ plays the role of m in the definition.
[2] Here, l and $m+n$ play the role of m and n, respectively, in the definition.
[3] Here, m and n play the role of l and m, respectively.

(ii) For each n, if we assume $P(n)$ is true: $(l+m)n = ln + mn$, then

$$(l+m)(n+1) = (l+m)n + (l+m) = (ln+mn) + (l+m)$$

by multiplication definition (ii) and substituting.

But now we may use the associative and commutative rules for addition to regroup and reorder the terms of the right side to get:

$$\begin{aligned}(ln+mn)+(l+m) &= ((ln+mn)+l)+m \\ &= (ln+(mn+l))+m \\ &= (ln+(l+mn))+m \\ &= ((ln+l)+mn)+m \\ &= (ln+l)+(mn+m)\end{aligned}$$

On the other hand:

$$l(n+1) = ln + l \text{ and } m(n+1) = mn + m$$

by multiplication definition (ii). So we see that $P(n+1)$ follows:

$$(l+m)(n+1) = l(n+1) + m(n+1)$$

and we conclude that the distributivity sentence $P(n)$ is true for **all** n.

We will leave the associative law of multiplication for the exercises, and prove instead the commutative law:

Multiplication is commutative: $mn = nm$ for all m and n.

Proof: Think of m as fixed and use proof by induction to prove:

$$\text{“}mn = nm\text{”}$$

(the commutativity sentences, which we'll call $P(n)$) for all values of n.

(i) $m \times 1 = m$, so we need to prove $1 \times m = m$. We'll do this by induction!

(i′) $1 \times 1 = 1$ by multiplication definition (i)

(ii′) For each n, $1 \times n = n$ implies that $1 \times (n+1) = 1 \times n + 1 = n + 1$ by multiplication definition (ii).

(i′) and (ii′) together prove $1 \times n = n$ for all n (including m, whatever m is), which is what we needed to prove (i). Now on to the proof of (ii).

(ii) For each n, $mn = nm$ implies that $m(n+1) = mn + m = nm + m$ by multiplication definition (ii) and substitution. But $nm + m = (n+1)m$ by the distributive law and (i) above, so $m(n+1) = (n+1)m$. By our proof by induction strategy we've proved $P(n)$ for all n, which is the commutative law.

This completes the basic arithmetic of the natural numbers. We've seen how induction was an important tool for making precise definitions and proofs.

Next, we'll do a couple of easy proofs by induction as further illustration of this powerful strategy. For this, we'll use one more definition:

Definition of subtraction of 1: For all natural numbers n except $n = 1$, $n - 1$ is defined to be the natural number immediately before n .

Proposition 1.1.1. *For all n, the nth odd number is $2n - 1$.*

Proof (by induction):

(i) The first odd number is 1, which is the number immediately before 2×1.

(ii) For each n, if the nth odd number is $2n-1$, then the $n+1$st odd number is $(2n - 1) + 2$ (odd numbers alternate with even numbers!), which is clearly(!) the number before $2n + 2 = 2(n + 1)$. That is, the $n + 1$st odd number is $2(n + 1) - 1$. End of proof!

Proposition 1.1.2. *For all n, the sum of the first n odd natural numbers is $n^2 = n \times n$.*

Proof (by induction):

(i) The sum of the first 1 odd numbers is $1 = 1^2$.

(ii) For each n, if the sum of the first n odd numbers is n^2, then the $n+1$st odd number is $2(n+1)-1 = 2n+1$ by Proposition 1.1.1 and so the sum of the first $n + 1$ odd numbers is:

$$n^2 + (2n + 1) = (n^2 + n) + (n + 1) = n(n + 1) + (n + 1) = (n + 1)^2$$

by the standard rules of arithmetic. End of proof.

Remark: One can prove many results of this type using induction. See the exercises and any elementary textbook in number theory.

Like subtraction, division of natural numbers is not usually defined. We'll fix this later with the rational numbers. However, we do have:

Division with remainders (long division): To each pair $m < k$, there is a quotient natural number q with the property that either:

(a) $k = mq$, and we say that m **divides** k (or m is a **factor** of k),

or else there is a remainder natural number $r < m$ such that:

(b) $k = mq + r$.

Proof: First of all, if $m = 1$ then m divides every number ($k = 1 \times k$) so we only need to worry about m's that are greater than 1. Fixing m, we will use proof by induction to prove division with remainders for all n and $k = m + n$ (the entire division with remainders sentence with $k = m+n$ will be our $P(n)$). The two cases complicate matters somewhat, but it all works out:

(i) $m + 1 = m \times 1 + 1$ (this is case (b) with $r = 1$). This proves $P(1)$.

(ii) For each n, if $P(n)$ is true for $k = m + n$, then either:

(a) $m + n = mq$, or or else:

(b) $m + n = mq + r$ and $r < m$

In (a), $m + (n + 1) = mq + 1$ which is case (b) of $P(n + 1)$ with remainder 1.

In (b), there are two possibilities to consider:

(a′) if $r + 1 = m$, then $m + (n + 1) = mq + m = m(q + 1)$
(This is case (a) of $P(n + 1)$) or

(b′) if $r + 1 < m$, then $m + (n + 1) = mq + (r + 1)$
(This is case (b) of $P(n + 1)$ with remainder $r + 1$).

Thus no matter which case occurs in division with remainders for $k = m+n$, we've proved that division with remainders is then true for $k = m + (n + 1)$. So division with remainders is true for **every** $k > m$ by induction.

Definition: A natural number p (other than 1) is **prime** if the only numbers that divide p are 1 and p itself.

The primes under 50: $2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47$.

Here's a different use of the well-ordered axiom.

Fundamental Theorem of Arithmetic: Every natural number except 1 factors as a product of prime numbers.

Proof: Let S be the set of natural numbers (other than 1) that **cannot** be factored as a product of prime numbers. By the well-ordered axiom S is either empty, or else it has a smallest element, which we'll call k, and which is either prime or not prime. But:

(a) Prime numbers do not belong to S, because they are factored!

(b) Numbers that are not prime (other than 1) can be written as $k = mq$ with both m and q smaller than k. So if k is the smallest element of S, then both m and q are not in S and therefore they can both be factored(!). But then k can be factored, too, so k couldn't have been in S in the first place!

Thus there cannot be a smallest element of S, so S is empty, which means that all natural numbers (other than 1) can be factored.

Remarks: The primes are the building blocks of the natural numbers much like the elements are the building blocks of the molecules. The factorization of a number into a product of primes is unique, as we shall see later in the course. If you are given a very large number, however, there is no known algorithm for factoring it quickly as a product of primes, even on a fast computer. This observation has been exploited for internet security. It is even difficult to tell quickly whether the number is prime or not, though there are some quick ways of telling "almost for sure" whether or not a given large number is prime.

We can now prove a very famous old result of Euclid:

Theorem (Euclid): There are infinitely many prime numbers.

Proof: Suppose we only know of finitely many. We could name them:

$$p_1, p_2, ..., p_n$$

and we could then build the following number:

$$N = p_1 \times p_2 \times ... \times p_n + 1$$

By the fundamental theorem of arithmetic, N factors as a product of primes, *but none of the known primes* $p_1, ..., p_n$ *divides* N. So the prime factors of N are new, and since there are always new primes, there must be infinitely many!

We finish the natural numbers with another useful result of Euclid:

Euclid's Algorithm: Start with natural numbers $m < n$. Perform long divisions according to the following algorithm whenever there is a remainder:

$$\begin{aligned} n &= mq_1 + r_1 \\ m &= r_1 q_2 + r_2 \\ r_1 &= r_2 q_3 + r_3 \\ r_2 &= r_3 q_4 + r_4 \end{aligned}$$

Eventually the algorithm terminates with no remainder: $r_k = r_{k+1} q_{k+2}$, and the last of the remainders, namely r_{k+1}, is the greatest common divisor (gcd) of m and n, i.e. the largest natural number that divides both. We will prove this later in the course, but for now, let's consider an:

Example: Find the gcd of $2^5 \times 3^3 \times 5^2 = 21600$ and $2^2 \times 3^3 \times 5^5 = 337500$.

We can do this with Euclid's algorithm:

$$\begin{aligned} 337500 &= 21600 \times 15 + 13500 \\ 21600 &= 13500 \times 1 + 8100 \\ 13500 &= 8100 \times 1 + 5400 \\ 8100 &= 5400 \times 1 + 2700 \\ 5400 &= 2700 \times 2 \end{aligned}$$

and then $r_k = 5400$, $r_{k+1} = 2700$ and $q_{k+2} = 2$, so 2700 is the gcd.

Of course, those of you who were paying attention in gradeschool also know: $2^2 \times 3^3 \times 5^2 = 2700$ is the gcd, and this is certainly an easier way to find the gcd of these two particular numbers than Euclid's algorithm. But there's a catch. This way of finding gcd's requires the **factorizations** of m and n. Euclid's algorithm doesn't require any factorizations, and indeed using Euclid's algorithm to find the gcd of two large numbers is very fast and easy (for computers), even when factoring the numbers is slow and hard.

1.1.1 Natural Number Exercises

1-1 Find the error in the following fake definition by induction of subtraction:

Fake Subtraction definition: Let m be a natural number other than 1.

(i) $f(1) = m - 1$ is the number before m.

(ii) $f(n+1) = m - (n+1)$ is the number before $f(n) = m - n$.

Therefore, $m - n$ is defined (and a natural number!) for all m and n.

1-2 Here's another arithmetic operation that can be defined by induction:

(a) Define the **exponential** m^n by induction.

(m is the "base" and n is the "exponent")

(b) Prove by induction that $m^{l+n} = m^l m^n$.

(c) Prove by induction that $(m^l)^n = m^{ln}$.

1-3 Define the **factorial** $n!$ by induction.

1-4 Prove by induction that addition of natural numbers is commutative.

1-5 Prove by induction that multiplication of natural numbers is associative.

1-6 (a) Prove by induction that the nth even number is $2n$.

(b) Prove by induction that the sum of the first n even numbers is $n \times (n+1)$.

1-7 Find (and count) all the prime numbers:

(a) between 1 and 100

(b) between 101 and 200

(c) between 201 and 300

(d) between 301 and 400

1-8 Factor each of the following numbers as a product of primes:

(a) 9699690

(b) 82861

(c) 10001

1-9 Find the gcd of 159477 and 241133 using Euclid's algorithm.

1-10 Induction also is important in set theory.

(a) Define the set $S_n = \{1, 2, ..., n\}$ by induction.

(b) Prove by induction that there are exactly 2^n subsets of S_n.

Hint: Exactly half the subsets of S_{n+1} contain the element $n + 1$.