Natural Numbers, Integers, and Rational Numbers (Following MacLane)

Abstract

We begin our rigorous development of number theory with **definitions** of $\mathbb{N}$ (the **natural numbers**), $\mathbb{Z}$ (the **integers**), and $\mathbb{Q}$ (the **rational numbers**). These definitions are complex, but they are the result of many and various observations about the way in which numbers arise. The first speculations about the various ways of *counting* and *listing* result in the **Peano axioms** for $\mathbb{N}$. This is our first example of the process of *abstraction*, and perhaps the most complex. Next, using **equivalence relations** and **equivalence classes** on $\mathbb{N}$ we *construct* $\mathbb{Z}$, and using equivalence relations and classes on $\mathbb{Z}$ we *construct* $\mathbb{Q}$. These constructions will be seen to be concrete realizations of our *axiomatizations* of $\mathbb{Z}$ and $\mathbb{Q}$, analogous to the Peano axioms for $\mathbb{N}$. In order to carry this out, we will of course first have to understand **relations** and **functions**, and therefore we beging with these. We leave the real numbers $\mathbb{R}$ for later, for these require rather different axioms and constructions.

1 Relations and Functions

Definition 1.1 *It is not necessary to have natural numbers defined ahead of the idea of* **ordered pair** (a, b)*, since* (a, b) *could be defined in a purely set-theoretical way, as follows (definition due to Kuratowski):*

$$(a, b) \stackrel{\text{def}}{=} \{\{a\},\ \{a, b\}\}$$

It should be clear that $(a, b) \neq (b, a)$*, as sets.* ■

Definition 1.2 *Let* X *and* Y *be sets. Their* **Cartesian product** $X \times Y$ *consists of ordered pairs* (x, y) *where* $x \in X$ *and* $y \in Y$*,*

$$X \times Y \stackrel{\text{def}}{=} \{(x, y) \mid x \in X,\ y \in Y\}$$

If $Y = X$*, we usually write* X^2 *instead of* $X \times X$*.* ■

Example 1.3 *We can view the xy-plane as the Cartesian product of* $\mathbb{R}$ *with itself,*

$$\mathbb{R}^2 = \mathbb{R} \times \mathbb{R} = \{(x, y) \mid x, y \in \mathbb{R}\}$$ ■

Definition 1.4 *If* X *and* Y *are any sets, then* any subset R *of their Cartesian product,*

$$R \subseteq X \times Y$$

is called a **binary relation** *on* X *and* Y*. If* $X = Y$*, we say* R *is a* relation on X*. Elements* (a, b) *in* R *are said to be* R*-related, and this is frequently denoted*

$$aRb$$

instead of $(a, b) \in R$*. Another common notation for relations is the tilde,* $\sim$*, but we will reserve this for a special type of relation, the equivalence relation.*

Every relation R *on* X *and* Y *has a* **domain** *and a* **range***,*

$$\mathcal{D}(R) \stackrel{\text{def}}{=} \{x \in X \mid \exists y \in Y,\ xRy\}$$
$$\mathcal{R}(R) \stackrel{\text{def}}{=} \{y \in Y \mid \exists x \in X,\ xRy\}$$

The set Y *is called the* **codomain** *of* R*. Thus,* $\mathcal{D}(R) \subseteq X$*, and* $\mathcal{R}(R) \subseteq Y$*.* ■

Remark 1.5 R *is* not necessarily *of the form* $\mathcal{D}(R) \times \mathcal{R}(R)$*, and in fact*

$$R \subseteq \mathcal{D}(R) \times \mathcal{R}(R) \subseteq X \times Y$$

We usually picture $\mathcal{D}(R) \times \mathcal{R}(R)$ *as a type of 'rectangle.'* ■

Here are some examples of relations:

Example 1.6 Inequalities, $\leq$ *and* $\geq$*, are binary relations on* $\mathbb{N}$*, called* **partial order** *relations. Strict inequalities* $<$ *and* $>$ *are also relations, but observe that* $n \leq n$ *but* $n \not< n$ *for any* $n \in \mathbb{N}$*.* ■

Example 1.7 Equality, $=$, *is a binary relation on any set* X. *Sometimes (e.g. in topology and other applications) equality is denoted* Δ, *which stands for* **diagonal**, *for it can be 'pictured' as the 'diagonal line'* $y = x$ *as in the* xy*-plane which is here* $X \times X$. ■

From this point of view, there are many relations:

Example 1.8 *Let* $R = \{(x, y) \mid x^2 + y^2 = 1\}$. *This is the unit circle* S^1, *but as a set explicitly stating that* x *and* y *are* R*-related by the equation* $x^2 + y^2 = 1$, *derived from the Pythagorean theorem.* ■

Remark 1.9 *The previous example illustrates a theme developed in* **algebraic geometry**: *the theme of* **zeros of polynomials**, *e.g. the zeros of* $p(x, y) = x^2 + y^2 - 1$. *These are also called* **algebraic curves**, *and as of today their study has evolved into the cutting edge of geometry.*

Definition 1.10 *Let* X *and* Y *are sets and* $A \subseteq X$, *then a* **function** *from* X *to* Y *is a binary relation* f *on* X *and* Y *with* **domain** $\mathcal{D}(f) = A$, *denoted here specially by*

$$f : A \to Y$$

instead of $R \subseteq X \times Y$, *and*

$$f(x) = y$$

instead of xRy *(for* $x \in A$ *and* f*-related* $y \in Y$*), and* <u>*all-importantly satisfying*</u>

$$f(x) = y \quad \textit{and} \quad f(x) = z \quad \Longrightarrow \quad y = z \tag{1.1}$$

for all $x \in X$ *and* $y, z \in Y$. *That is,* <u>*only one* y*-value to each* x</u>, *colloquially known as the* **vertical line test**, *which tests whether a random relation* R *is a function by passing a vertical line through* R: *it should intersect* R *in at most one point at a time. One more time, now:* A *is the domain,* Y *the codomain, but now let us also include the* **range**, *or* **image**, *of* f

$$f(A) = \{y \in Y \mid \exists x \in A \textit{ such that } y = f(x)\} \tag{1.2}$$

The **graph** *of* f

$$\begin{aligned} \text{graph}(f) &\overset{\text{def}}{=} \{(x, y) \mid x \in A,\ y \in f(A)\} && (1.3)\\ &= \{(x, f(x)) \mid x \in A\} && (1.4) \end{aligned}$$

is thus identical with the definition of f *as a set,* $f = \text{graph}(f)$. ■

Let us now consider **special types of relations**. We need these to get to equivalence relations, which is the right tool for constructing $\mathbb{Z}$ and $\mathbb{Q}$. Here, $X = Y$ and $R \subseteq X^2$.

- R is **reflexive** if aRa for all $a \in A$ (e.g. $\leq$ and $=$ on $\mathbb{N}$)
- R is **irreflexive** if $a \not R a$ for all $a \in A$ (e.g. $<$ on $\mathbb{N}$)
- R is **symmetric** if $aRb \implies bRa$ for all $a, b \in A$ (e.g. $=$ on $\mathbb{N}$)
- R is **asymmetric** if $aRb \implies b \not R a$ for all $a, b \in A$ (e.g. $<$ on $\mathbb{N}$)
- R is **antisymmetric** if aRb and $bRa \implies a = b$ for all $a, b \in A$ (e.g. $\leq$ on $\mathbb{N}$)
- R is **transitive** if aRb and $bRc \implies aRc$ for all $a, b, c \in A$ (e.g. $\leq$, $<$ and $=$ on $\mathbb{N}$)

2 Equivalence Relations

Definition 2.1 *An* **equivalence relation** *on a set X is a binary relation which is*

(1) reflexive
(2) symmetric
(3) transitive

Its definition reflects our Platonic desire to define sameness of properties in different individuals. *We simply group those individual objects (now residing in a definite set X) exhibiting the 'same' property into a subset and call this fact an equivalence relation between them. This will be argued explicitly below.* ■

Example 2.2 *Congruence $\cong$ and similarity $\sim$ of triangles, or indeed of any 'figures,' in the Euclidean plane $\mathbb{R}^2$ are two properties that give rise to two equivalence relations: triangles having the same length sides are congruent but* not necessarily equal, *and likewise triangles that possess the same angles are similar but* not necessarily equal. *Clearly a single triangle is both congruent and similar to itself, so both $\cong$ and $\sim$ are relexive. Moreover, two triangles A and B have $A \cong B \implies B \cong A$ and $A \sim B \implies B \sim A$,*

so both $\cong$ and $\sim$ are symmetric. Finally three triangles A, B and C have $A \cong B$ and $B \cong C \implies A \cong C$, and likewise with $\sim$, so both $\cong$ and $\sim$ are transitive. These three traits capture an equality among triangles that is broader than strict identity, and this is what we want our equivalence relation to do in general. ■

Definition 2.3 *If $\sim$ is an equivalence relation on X and $a \in X$, then the* **equivalence class** *of a is the set of all $x \in X$ which satisfy $x \sim a$, denoted*

$$[a] = \{x \in X \mid x \sim a\}$$

The set of all equivalence classes on X *is denoted $X/\sim$, and is called the* **quotient set on X by $\sim$**. *The function $\pi : X \to X/\sim$ given by $\pi(x) = [x]$ is called the* **canonical projection**, *or the* **quotient map**, *of $\sim$.* ■

Definition 2.4 *If X is a set, and $\mathscr{P}$ is a collection of subsets of X, that is $\mathscr{P} \subseteq \mathcal{P}(X)$, satisfying:*

(1) $\varnothing \notin \mathscr{P}$.

(2) $\bigcup_{A\in\mathscr{P}} A = X$ *where* $\bigcup_{A\in\mathscr{P}} A \stackrel{\text{def}}{=} \{a \in A \mid A \in \mathscr{P}\}$ *is the* **union** *of all the A in $\mathscr{P}$.*

(3) *All* distinct $A, B \in \mathscr{P}$ *are* **disjoint**, $A \cap B = \varnothing$.

then we call $\mathscr{P}$ a **partition** *of X. In plain English, a partition is a collection of disjoint nonempty subsets of X whose union is X. Incidentally, we can combine 'union' and 'disjoint' into one entity, the* **disjoint union** *of the the $A \in \mathscr{P}$ by using a* square *cup:*

$$\bigsqcup_{A\in\mathscr{P}} A = X$$

This combines (2) *and* (3) *into one statement. Alternative notations for union $\bigcup_{A\in\mathscr{P}} A$ and disjoint union $\bigsqcup_{A\in\mathscr{P}} A$ are*

$$\bigcup \mathscr{P} \equiv \bigcup_{A\in\mathscr{P}} A \qquad \textit{and} \qquad \bigsqcup \mathscr{P} \equiv \bigsqcup_{A\in\mathscr{P}} A$$ ■

Lemma 2.5 *If X is a set and $\sim$ is an equivalence relation on X, then two equivalence classes $[a]$ and $[b]$ of X are either disjoint or equal. Consequently, $X/\sim$ is a partition of X*

Proof: Suppose $[a] \cap [b] \neq \varnothing$ and let $x \in [a] \cap [b]$. Then, $x \sim a$ and $x \sim b$. By symmetry, $a \sim x$, and by transitivity $a \sim x$ and $x \sim b \implies a \sim b$. Also, $\forall y \in [a],\ y \sim a$, and by transitivity again $y \sim a$ and $a \sim b \implies y \sim b \implies y \in [b]$, or $[a] \subseteq [b]$. Similarly, $\forall z \in [b],\ z \sim a \implies [b] \subseteq [a]$, and so $[a] = [b]$. But then the set of all equivalence classes determined by $\sim$, namely the quotient set $X/\sim$, is a partition of X, since all it's elements are pairwise disjoint and their union is X. ■

Theorem 2.6 *If X is a set and $\mathscr{P}$ is a partition of X, then there is exactly one equivalence relation on X from which it is derived.*

Proof: Define a binary relation $\sim$ on X by setting $x \sim y$ if $x, y \in A$ for some $A \in \mathscr{P}$. Now, $\sim$ is obviously symmetric, reflexive and transitive, and the relation applies to all elements of X, since $\bigsqcup \mathscr{P} = X$ and all elements of $\mathscr{P}$ are pairwise disjoint. Now, suppose there were two equivalence relations $\sim_1$ and $\sim_2$ on X with the above properties. Then, $\forall a \in X$ let $[a]$ and $[a']$ be the equivalence classes determined by a relative to $\sim_1$ and $\sim_2$, respectively. Consequently, $[a], [a'] \in \mathscr{P}$ must equal the unique element $A \in \mathscr{P}$ containing a (by the above paragraph), i.e. $\{x \in A \mid x \sim_1 a\} = [a] = A = [a'] = \{x \in A | x \sim_2 a\}$, which implies $\sim_1 = \sim_2$. Thus, $\sim$ is the unique equivalence relation from which $\mathscr{P}$ is derived. ■

3 Natural Numbers

3.1 The Peano Axioms for $\mathbb{N}$

As MacLane says (p. 42), long human experience with counting and listing has resulted in the distillation of the ***rules of reckoning*** to a short ***formal*** list:

(1) $n + 0 = n$ (0 *is the* ***additive identity***)

(2) $n + m = m + n$ (***commutativity*** *of* $+$)

(3) $(n + m) + p = n + (m + p)$ (***associativity*** *of* $+$)

(4) $1n = n = n1$ (1 *is the* ***multiplicative identity***)

(5) $mn = nm$ (***commutativity*** *of* $\cdot$)

(6) $(mn)p = m(np)$ (***associativity*** *of* $\cdot$)

(7) $p(m + n) = pm + pn$ (***distributivity*** *of* $\cdot$ *over* $+$)

for all $m, n, p \in \mathbb{N}$.

> *'[T]hese (long-established) rules are inviolate: If it doesn't turn out as they specify, I know that I have made a mistake somewhere. This is the merit of a formal rule: Once firmly established, it can be applied mechanically and is an infallible guide.'* — MacLane, p. 42

From Aristotle's point of view, axioms are the foundation of our logical analysis of an idea (a *Platonic* idea). The modern point of view largely agrees with this, except it doesn't interpret axioms in the same way as Aristotle. To Aristotle, axioms were *self-evident, directly discernible in experience*. The modern view is more cosmopolitan: the seven axioms above are *one* of possibly many beginnings. Turns out the ***Peano axioms*** or ***postulates*** are more flexible and provide a shorter list. As MacLane observes:

> *'...there is no answer to the question: What is a natural number? There are various alternative concrete descriptions, depending on the sort of counting intended or on the prior assumptions. In each case, the description provides "numbers" which do satisfy the Peano postulates. Hence we conclude that one does not define what a natural number "is", by itself. Instead, one defines the system of all natural numbers, with successor operation. Then $\mathbb{N}$ is any such system which satisfies the Peano postulates. This means that there are many such systems within set theory—but that they are all isomorphic...*
>
> *Note that the postulates themselves are by no means unique; for example, the Peano postulates may be replaced by the recursion theorem as an axiom. Here, as in other axiomatic descriptions of mathematical objects, there are a variety of choices for lists of axioms. Number theory, like other subjects in Mathematics, is not the study of a unique model nor yet the examination of a unique axiomatic system—it is rather a study of the* form *exemplified by the various models and specified in the axioms.'* — MacLane, p. 42

Not only can we reduce the original axiom list to five, and get double the mileage, by switching to the Peano axioms and then *deriving* the longer list as a consequence, but we notice that we could equally have chosen the recursion theorem as an axiom. This flexibility is a neat modern flourish, but significant, because this smaller list—or it's recursion equivalent—actually has the ***principle of mathematical induction*** built-in. That *wasn't* in the original list of seven properties, so by this ruse we have thrown in a mega bonus property. This is the property that allows us to compute π by inscribed polygons, for example, and it is also the property that lets us use the Intermediate Value Theorem iteratively to construct a root-finding algorithm for equations. And, perhaps most important of all, mathematical induction is a basic workhorse proof method in every area of math, without which many theorems would need re-proving some other way, were that always possible.

Another point needs consideration. As MacLane observes, 'number theory, like other subjects in Mathematics, *is not the study of a unique model* nor yet the examination of a unique axiomatic system—it is rather a study

of the form *exemplified by the various models* and specified in the axioms.' What are these 'models' he's referring to? Aside from the axioms, which, as MacLane says, specify the *form* of any 'model,' we need a **'model'**: a concrete realization, in some defined setting. This can be formalized as follows: axioms specify the form of any concrete realization, or interpretation, of the axioms in a **structure**: at the very least in terms of sets, but usually endowed with extra bells and whistles, like algebraic operations, say. It turns out to be possible to *construct* $\mathbb{N}$ as a set with all the right trappings so that it satisfies Peano's axioms. This is done in some detail in Enderton's book, 'Elements of Set Theory,' Chapter 4. We will give a much shorter description of this formalism below.

Definition 3.1 (Peano Axioms for $\mathbb{N}$) *The* **Peano axioms** *for the natural numbers $\mathbb{N}$ go as follows: Without worrying overmuch about what '0' '1,' and 'successor of n'* are *in any Platonic sense—as we said, these are axioms, not models or realizations—we require*

(1) $0 \in \mathbb{N}$

(2) $n \in \mathbb{N} \implies s(n) \stackrel{\text{def}}{=} n+1 \in \mathbb{N}$ *(***successor of** *n)*

(3) $0 \neq s(n)$ *for any* $n \in \mathbb{N}$

(4) $s(m) = s(n) \implies m = n$ *for all* $m, n \in \mathbb{N}$

(5) *Let P be a property (hopefully of all $\mathbb{N}$). If 0 has P and whenever n has P so does $s(n)$, then all $n \in \mathbb{N}$ have P. (***Principle of Mathematical Induction***, non-set-theoretic)* ■

3.2 The Recursion Theorem

Consider now the notion of **recursion**.

Example 3.2 *Inscribe a regular hexagon in the unit circle, and from elementary geometry you find that its sidelength s_1 is 1, and therefore its perimeter is 6. This is our first approximation of 2π, the circumference of the circle, and it gives 3 as a first approximation of π. Let $a_1 = 3$, and define the side-length sequence s_n recursively as follows: Once we know s_1, we double the number of sides to inscribe a regular 12-gon, and again from elementary geometry deduce that the new sidelength is*

$$s_2 = \sqrt{2 - 2\sqrt{1 - \frac{s_1^2}{4}}}$$

In fact, once this trick is observed, it can be applied to get s_3 in terms of s_2, s_4 in terms of s_3, etc. In general, once we know s_n, we can get

$$s_{n+1} = \sqrt{2 - 2\sqrt{1 - \frac{s_n^2}{4}}} \tag{3.1}$$

Now, back to our approximating sequence a_n for π. Our regular polygons have, first, $6 = 3 \cdot 2$ sides ($n = 1$ here), then $12 = 3 \cdot 2^2$ sides ($n = 2$ here), etc. Half of this is our approximation to π,

$$a_n = 3 \cdot 2^{n-1} \cdot s_n$$

■

The example above illustrates the technique. We started with

$$s_1 = 1$$

and defined

$$s_{n+1} = \sqrt{2 - 2\sqrt{1 - \frac{s_n^2}{4}}}$$

How should we interpret this? To connect it to the following Recursion Theorem, let us introduce some notation. If we define the functions

$$f : \mathbb{N} \to \mathbb{R}, \quad f(n) \stackrel{\text{def}}{=} s_n$$

$$g : \mathbb{R} \to \mathbb{R}, \quad g(x) = \sqrt{2 - 2\sqrt{1 - \frac{x^2}{4}}}$$

then the above two equations may be rephrased

$$f(1) \stackrel{\text{def}}{=} 1$$
$$f(s(n)) \stackrel{\text{def}}{=} g(f(n))$$

We have thus defined f recursively, using the ***recurrence relation*** g. We defined it for $n = 1$, and used $f(1)$ and g to get $f(2)$, then used $f(2)$ and g to get $f(3)$, etc. Since recursion is programmable/implementable, I used it to construct a MATLAB 'for' loop to run the computation ten steps for me:

```
s = 1;
P = 3*s;
fprintf('s = %.10f,         P = %.10g \n',s,P)

for i = 1:10
    s = sqrt(2-2*sqrt(1-s^2/4));
    P = 3*(2^i)*s;
    fprintf('s = %.10f,         P = %.10g \n',s,P)
end
```

Remark 3.3 *Notice that $f : \mathbb{N} \to \mathbb{R}$ gives the values of a sequence, $f(n) = s_n$. This is a general fact. Any sequence $(a_n)_{n\in\mathbb{N}}$ can be thought of as a function $f : \mathbb{N} \to X$, where $f(n) = a_n \in X$. The set X may be a probability space, a manifold, a function space, or any other set.* ■

We thus take the Recursion Theorem as a consequence of the Peano axioms, but as MacLane observes, this logical dependency could be reversed, since the recursion theorem is logically equivalent to the Peano axioms:

Theorem 3.4 (Recursion Theorem) *Let X be any set and let $a \in X$ be a fixed element. Given any function $g : X \to X$ we can* construct*, or more specifically* recursively define*, a function $f : \mathbb{N} \to X$ using a, g and the Peano axioms, by*

$$f(0) \stackrel{\text{def}}{=} a \tag{3.2}$$
$$f(s(n)) \stackrel{\text{def}}{=} g(f(n)) \tag{3.3}$$

Remark 3.5 *In my example above, I started with $f(1) = 1$, instead of $f(0) = 1$, but this was only a cosmetic modification. I could just as easily have written $s_0 = 1$ instead of $s_1 = 1$, and made $f(0) = 1$.* ■

Remark 3.6 *As MacLane remarks (p. 45), 'A proof of this theorem uses axiom (v') and must depend upon the set-theoretic definition of "function."' Axiom (v') is an alternate version of our Peano axiom (5):*

(5) *If $S \subseteq \mathbb{N}$ is a set of natural numbers containing 0 and if every $n \in S$ has its successor $s(n) \in S$, then $\mathbb{N} \subseteq S$. (***Principle of Mathematical Induction***, set-theoretic)*

Since $S \subseteq \mathbb{N}$ and $\mathbb{N} \subseteq S$, we conclude that $S = \mathbb{N}$. Thus, if we use sets *S instead of properties P, we can get our proof, as in MacLane:* ■

Proof: To use this axiom (5) we need our set S. Let

$$\mathbf{0} = \{0\},\ \mathbf{1} = \{0, 1\},\ \ldots\ , \mathbf{n} = \{0, 1, \ldots, n\}$$

and let P be the property of n

$$P = \text{'There is a unique function } f_n : \mathbf{n} \to X \text{ satisfying (3.2)-(3.3).'}$$

We observe that, for $n = 0$, the theorem's hypothesis gives us $f(0) = a$ in (3.2), so we call this function $f_0 : \mathbf{0} \to X$. For $n = 1$ we'll have, by (3.3), that we can *define* $f(1) = f(s(0)) = g(f_0(0))$, and we call this $f_1 : \mathbf{1} \to X$,

$$\begin{aligned} f_1(0) &= f_0(0) = a \\ f_1(1) &= g(f_0(0)) = g(a) \end{aligned}$$

Next, we define $f(2) = f(s(1)) = g(f(1)) = g(g(a))$ using (3.2)-(3.3) and our previous result. Call this $f_2 : \mathbf{2} \to X$,

$$\begin{aligned} f_2(0) &= f_1(0) = f_0(0) = a \\ f_2(1) &= f_1(1) = g(f_1(0)) = g(a) \\ f_2(2) &= g(f_1(1)) \end{aligned}$$

the last of which, incidentally, equals $g(g(a))$ or $g^2(a)$ for short. Continuing in this way, once we have $f_n : \mathbf{n} \to X$ defined, we let f_{n+1} take on the same values on 0, 1, ..., n, and define $f_{n+1}(n+1) = g(f_n(n))$ using (3.3). The list

of functions $f_0, \ldots, f_{n+1}$ fit together, in the sense that f_{n+1} uses the other f_i's previously defined and only adds the next value. Let

$$S = \bigcup_{n \in \mathbb{N}} \mathbf{n}$$

and note that S satisfies the desired properties if we define $f = \bigcup_{n \in \mathbb{N}} f_n$, that is if we define $f(n) = f_n(n) = g(f_{n-1}(n-1))$, and this equals $g^n(a)$, incidentally. ■

All of this formalism is designed to make you feel better about plugging $g^n(a)$ back into g to get $g^{n+1}(a)$, because this is how we're defining $f(n+1)$! At least 'better' in the sense of having something to say for yourself when someone on the street asks you how you know how $f(n)$ may be defined in advance for infinitely many n. Tell them it's to talk to Peano.

3.3 The Set-Theoretic Construction of the Natural Numbers

To use sets as building blocks for a concrete realization of $\mathbb{N}$, that is to construct a set-theoretic model, requires, as with everything in modern math, a choice of set theory. The current convention is Zermelo-Fraenkel with Choice (ZFC), whose nuances I leave to another course in the foundations of math. For us, the important axiom in ZFC for the construction of $\mathbb{N}$ is Zermelo's Axiom of Infinity:

(Infinity) (Zermelo, 1908) There exists an infinite set, specifically an inductive set containing $\varnothing$.

An inductive set is exactly the sort of set you need to suppose the successor function is defined on it. Nevermind about the details. The point is the following: we merely assume we have a set X on which a successor function can be defined (if somebody asks you how you got hold of this successor function, tell them you just assumed it was there for the taking, and that if they don't like that then they shouldn't ask for a model of $\mathbb{N}$. That should quiet them.).

Definition 3.7 (Set-Theoretic Construction of $\mathbb{N}$) *We want to define the natural numbers using only the axiom of infinity, the empty set $\varnothing$ and curly brackets $\{\cdot\}$, and it should be ordered and have all of the known arithmetic properties known from elementary math. Since we are using the axiom*

of infinity, we have an infinite set N and a successor function on it, which we define by

$$S : N \to N$$
$$S(n) \stackrel{\text{def}}{=} n \cup \{n\}$$

Here, n must be a set, of course, so it makes sense to union it with another set. Moreover, we require $S(n)$ to contain n for all $n \in N$,

$$\forall n \in N,\ n \in S(n)$$

The order relation $<$ on N then simply becomes $\in$,

$$n < m \iff^{\text{def}} n \in m, \quad \text{for all } n, m \in N$$

Moreover, keeping those Peano axioms in mind, we don't want 0 to be the successor of any natural number, so we require that

$$\forall n \in N, 0 \neq S(n)$$

A way to do all of this is to identify the symbol 0 with $\varnothing$,

$$0 \stackrel{\text{def}}{=} \varnothing$$

Also, we would like each natural number n to have n elements as a set, in such a way as to mesh with our definition of $<$. In particular, we demand that $n < n + 1$, i.e.

$$n \in n + 1 \tag{3.4}$$

Then, too, whenever $m < n$, i.e. $m \in n$, we must have $m \in n + 1$, or $m < n + 1$ by the transitivity of $< \iff \in$. We must accordingly require

$$n \subseteq n + 1 \tag{3.5}$$

So since $n \in n + 1$ and $n \subseteq n + 1$ by (3.4) and (3.5), we see that we must define S and $n + 1$ simultaneously by

$$S(n) \stackrel{\text{def}}{=} n \cup \{n\} \stackrel{\text{def}}{=} n + 1$$ ■

For example, since

$$0 \stackrel{\text{def}}{=} \varnothing$$

1 and 2 are defined by

$$\begin{aligned} 1 &\stackrel{\text{def}}{=} 0+1 \\ &= S(0) \\ &= S(\varnothing) \\ &= \varnothing \cup \{\varnothing\} \\ &= \{\varnothing\} \end{aligned} \qquad \begin{aligned} 2 &\stackrel{\text{def}}{=} 1+1 \\ &= S(1) \\ &= 1 \cup \{1\} \\ &= \{\varnothing\} \cup \{\{\varnothing\}\} \\ &= \{\varnothing, \{\varnothing\}\} \end{aligned}$$

and so on. Using the symbols for the sets, this reduces to

$$\begin{aligned} 0 &= \varnothing \\ 1 &= \{0\} \\ 2 &= \{0,1\} \\ &\vdots \\ n &= \{0,1,\ldots,n-1\} \end{aligned}$$

Theorem 3.8 (Uniqueness of $\mathbb{N}$) *There exists exactly one set $\mathbb{N}$ satisfying*

(1) $\varnothing \in \mathbb{N}$

(2) $n \in \mathbb{N} \implies S(n) \in \mathbb{N}$

(3) *If K is any set satisfying* (1) *and* (2)*, then $\mathbb{N} \subseteq K$.*

Proof: By the axiom of infinity there exists at least one set X satisfying (1) and (2). Let

$$\mathcal{F} = \{Y \in \mathcal{P}(X) \mid \varnothing \in Y \text{ and } x \in Y \implies S(x) \in Y\}$$

and

$$\mathbb{N} = \bigcap \mathcal{F}$$

Then $\mathbb{N}$ satisfies (1) and (2) because $\varnothing \in Y$ for all $Y \in \mathcal{F}$, so that $\varnothing \in \mathbb{N}$, and $x \in \mathbb{N} \implies x \in Y$ for all $Y \in \mathcal{F} \implies S(x) \in Y$ for all $Y \in \mathcal{F} \implies S(x) \in \mathbb{N}$. Now, if K is any set satisfying (1) and (2), then $X \cap K \in \mathcal{F}$, so that $\mathbb{N} = \bigcap \mathcal{F} \subseteq X \cap K \subseteq K$. Consequently, $\mathbb{N}$ is unique, for if $\mathbb{N}'$ is any other set satisfying (1)-(3), then $\mathbb{N}' \subseteq \mathbb{N}$ and $\mathbb{N} \subseteq \mathbb{N}'$ by (3), so that $\mathbb{N} = \mathbb{N}'$. ■

Remark 3.9 *Of course, if $K \subseteq \mathbb{N}$, then $K = \mathbb{N}$, which is the* **induction property** *of $\mathbb{N}$, and which demonstrates that $\mathbb{N}$ satisfies the 5th Peano Axiom. Of course, $\mathbb{N}$ satisfies axioms 1-3 by design, and it satisfies axiom 4 by the next theorem.* ■

Theorem 3.10 *For all $m, n, p \in \mathbb{N}$, as defined above, we have the following relations:*

(1) $n \subseteq n+1$ *and* $n \in n+1$

(2) $m \in n \implies m \in \mathbb{N}$

(3) $m \in n \implies m+1 \subseteq n$

(4) $m \in n$ *and* $n \in p \implies m \in p$

(5) $n \notin n$

(6) $n+1 = m+1 \implies m = n$

(7) $m \subseteq n$ *iff* $m = n$ *or* $m \in n$

(8) $m \subseteq n$ *or* $n \subseteq m$

Proof: (1) Since $n+1 = n \cup \{n\}$, we have both $n \subseteq n+1$ and $n \in n+1$.

(2) Let $K = \{n \in \mathbb{N} \mid m \in n \implies m \in \mathbb{N}\}$. Then $0 = \varnothing \in K$ trivially, since there is no $x \in 0$. Now, if $n \in K$, let $m \in n+1 = n \cup \{n\}$. Then, either $m \in n$ or $m = n$. In the first case we have by the definition of K that $m \in \mathbb{N}$, which means $n+1 \in K$ by the definition of K, since $m \in n+1$. In the second case we have that $m = n \in \mathbb{N}$, so again $n+1 \in K$. Either way, $n \in K \implies n+1 \in K$, and since $K \subseteq \mathbb{N}$, we have by the induction property of $\mathbb{N}$ that $K = \mathbb{N}$, which means for all $m \in n \implies m \in \mathbb{N}$ for all $n \in \mathbb{N}$.

(3) Let $K = \{n \in \mathbb{N} \mid m \in n \implies m+1 \subseteq n\}$. Obviously $0 = \varnothing \in K$ trivially, while if $n \in K$, consider $n+1$ and any $m \in n+1$. Either $m \in n$ or else $m = n$. In the latter case we have that $m+1 = n+1 \subseteq n+1$, while in the former we have that $m+1 \subseteq n \subseteq n+1$, so that $m+1 \subseteq n+1$. Thus, either way we have that $n+1 \in K$, so by the induction property of $\mathbb{N}$ we have that $K \subseteq \mathbb{N}$ and hence $K = \mathbb{N}$, or $m \in n \implies m+1 \subseteq n$ for all $m, n \in \mathbb{N}$.

(4) Suppose $m \in n$ and $n \in p$. Since $n \subseteq n+1$ we have that $m \in n+1$, and since by (3) we have that $n \in p$ implies that $n+1 \subseteq p$, we conclude that $m \in n+1 \subseteq p$, or $m \in p$.

(5) Let $K = \{n \in \mathbb{N}_0 \mid n \notin n\}$. Clearly $\varnothing \in K$, while if $n \in K$, then $n \notin n$. Suppose $n+1 \in n+1$. Then $n \cup \{n\} \in n \cup \{n\}$. Since $n \notin n$, we cannot have $n \cup \{n\} \in n$, so we must have $n \cup \{n\} \in \{n\}$. But then $n = n \cup \{n\}$, which is impossible since for that to be true we would need to have $\{n\} = \varnothing$. Hence $n+1 \notin n+1$, and so $K \subseteq \mathbb{N}_0$, whence $K = \mathbb{N}_0$, or $n \notin n$ for all $n \in \mathbb{N}$.

(6) If $n+1 = m+1$, then $n \cup \{n\} = m \cup \{m\}$, and suppose $n \neq m$. Then if by the axiom of extensionality, if $x \in n \cup \{n\}$, then $x \in m \cup \{m\}$. If $x = n$, then $n \in m \cup \{m\}$, and since by assumption $n \neq m$, we can't have $n \in \{m\}$, so we must have $n \in m$. But then we must also have $m \in n$ by the same reasoning, which is impossible because by (4) we'd have $m \in m$ and $n \in n$, which contradicts (5). Hence we must have $n = m$.

(7) If $m = n$ then $m \subseteq n$, while if $m \in n$ then by (1) and (3) we have $m \subseteq m+1 \subseteq n$, so $m \subseteq n$. Conversely, if $m \in \mathbb{N}_0$, let $K = \{n \in \mathbb{N}_0 \mid m \subseteq n \implies m \in n \text{ or } m = n\}$. Clearly $0 = \varnothing \in K$ trivially, while if $n \in K$, then choose $m \subseteq n \cup \{n\} = n+1$. If $m \not\subseteq n$, then there is some $x \in m \cap (n \cup \{n\}) \backslash n = \{n\}$, so that $n \in m$. But then by (2) we have $n+1 \subseteq m$, which combined with $m \subseteq n+1$ implies that $m = n+1$, so that $n+1 \in K$. If $m \subseteq n$, then $n \in K$ we have $m \in n$ or $m = n$: if $m \in n \subseteq n+1$ so $m \in n+1$, which means $n+1 \in K$, while if $m = n$ then clearly $m = n \in \{n\} \subseteq n \cup \{n\} = n+1$, or $m \in n+1$, and $n+1 \in K$. Thus in all cases $n \in K \implies n+1 \in K$. Consequently $\mathbb{N}_0 \subseteq K \subseteq \mathbb{N}_0$, or $K = \mathbb{N}_0$.

(8) Define the set $\{n \in \mathbb{N}_0 \mid n \notin m \implies n \subseteq m\}$. Of course $0 = \varnothing \in K$ trivially since $\varnothing \subseteq m$ for all $m \in \mathbb{N}$. Now, if $n \in K$, then let $m \in \mathbb{N}_0$ and suppose that $m \notin n+1 = n \cup \{n\}$. Then $m \neq n$ and $m \notin n$, which means that since $n \in K$ we must have $n \subseteq m$. But since $m \neq n$ we must have by (7) that $n \in m$, and by (3) that $n+1 \subseteq m$. Consequently $n+1 \in K$, and by the induction principle we have $K = \mathbb{N}_0$. To finish the proof note that if $n, m \in \mathbb{N}$ and $n \subseteq m$ then we're done. So assume that $m \not\subseteq n$. By (7) we have $m \notin n$, and since $m \in \mathbb{N}_0 = K$ we conclude that $m \subseteq n$. ■

Remark 3.11 *We will henceforth revert to the usual method of "proving by mathematical induction", namely omitting mention of the set K, as we did above, in order to keep with the conventions of the mathematicians. We will simply demonstrate that a given property holds for elements of a set indexed by $n = 0$ or $n = 1$ and then show that whenever the case for $n = k$ holds*

the case for $n = k + 1$ holds as well, concluding from this that the property holds for all elements indexed by $\mathbb{N}$. ■

Theorem 3.12 (Basic Arithmetic Properties of $\mathbb{N}$) *If we define binary operations of addition $+$ and multiplication $\cdot$ inductively on $\mathbb{N}$ by*

$$\begin{array}{lll} n + 0 = n & \textit{and} & n + (m + 1) = (n + m) + 1 \\ n0 = 0 & \textit{and} & n(m + 1) = (nm) + n \\ & & (m + 1)n = (mn) + n \\ n^0 = 1 & \textit{and} & n^{m+1} = n^m n \end{array}$$

for all $m, n \in \mathbb{N}$, then these operations satisfy the usual arithmetic properties: for all $m, n, p \in \mathbb{N}$, we have

(1) $m + (n + p) = (m + n) + p$ *(associativity of addition)*

(2) $m + n = n + p$ *(commutativity of addition)*

(3) $\exists 0 \in \mathbb{N}$ *such that* $n + 0 = n$ *(additive identity)*

(4) $m(np) = (mn)p$ *(associativity of multiplication)*

(5) $mn = nm$ *(commutativity of multiplication)*

(6) $\exists 1 \in \mathbb{N}$ *such that* $n1 = n$ *(multiplicative identity)*

(7) $m(n + p) = mn + mp$ *and* $(n + p)m = nm + pm$ *(distributivity of $\cdot$ over $+$)*

(8) $n + m = n + p \implies m = p$ *(cancellation law for $+$)*

(9) $nm = np \implies m = p$ *if* $n \neq 0$ *(cancellation law for $\cdot$)*

(10) $n + m \leq n + p \iff m \leq p$ *(cancellation law for $+$ and $\leq$)*

(11) $nm \leq np \iff m \leq p$ *if* $n \neq 0$ *(cancellation law for $\cdot$ and $\leq$)*

(12) $n^{m+k} = n^m n^k$ *(exponent law)*

Proof: We prove these in the order (3), (1), (2), (7), (6), (4), (5), (8), (9), (10), (11), (12): First, (3) follows from our definition of $+$.

(1) From (3) we have for all $m, n, p \in \mathbb{N}$ that

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

and if $(m+n)+p = (m+n)+p$, then

$$\begin{aligned}(m+n)+p+1 &= ((m+n)+p)+1 \\ &= (m+(n+p))+1 \\ &= m+((n+p)+1) \\ &= m+(n+(p+1))\end{aligned}$$

where the first, third and fourth equality follow from our definitin of addition, and the second by the induction hypothesis. Consequently associativity holds true for all $m, n, p \in \mathbb{N}$.

(3) First, note that $0 = 0+0$ by the definition of $+$. Now, suppose that $0+n = n$ for $n \in \mathbb{N}$. Then,

$$0+(n+1) = (0+n)+1 = n+1$$

so we have that $0+n = n$ for all $n \in \mathbb{N}$ as well. Thus, we have that $0+n = n+0$ for all $n \in \mathbb{N}$. Next, we prove the case $n+1 = 1+n$: we have $0+1 = 1+0$ by the above argument. If $n \in \mathbb{N}$ satisfies $n+1 = 1+n$, then by commutativity we have

$$(n+1)+1 = (1+n)+1 = 1+(n+1)$$

Consequently we have for all $n \in \mathbb{N}$ that $1+n = n+1$ by induction. Now, if $m, n \in \mathbb{N}$ satisfy $m+n = n+m$, then by the above argument and commutativity we have

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

Thus, by the induction principle we have that $m+n = n+m$ for all $m, n \in \mathbb{N}$.

(7) Clearly for all $n, p \in \mathbb{N}$ we have $0(n+p) = 0 = 0+0 = 0n+0p$. Now, if $m, n, p \in \mathbb{N}$ satisfy $m(n+p) = mn+mp$, then by our definition of

$n(p+1) = (np) + n$ and by associativity we have

$$\begin{aligned} m(n + (p+1)) &= m((n+p)+1) \\ &= (m(n+p)) + m \\ &= (mn + mp) + m \\ &= mn + (mp + m) \\ &= mn + m(p+1) \end{aligned}$$

Thus by induction we have that $m(n+p) = mn + mp$ for all $m, n, p \in \mathbb{N}$.

(4)-(6) and (8)-(12) follow similarly and are left as exercises! ■

Remark 3.13 *We have shown that* $(\mathbb{N}, \leq, +, \cdot, 0, 1)$ *is a* **structure** *satisfying the Peano Axioms which can be completely embedded in the sets of ZFC set theory.* ■

4 Integers

If we try to define a **subtraction** operation "$-$" on $\mathbb{N}$, we immediately notice that a given natural number k could be represented as the difference of two other natural numbers m and n, such as $5 = 7 - 2$. But *this representation is not unique*, since $5 = 7 - 2 = 23 - 18 = \cdots$. However, notice that

$$7 - 2 = 23 - 18 \iff 7 + 18 = 23 + 2$$

where the latter expression makes no use of any "$-$" operation. This gives us a way to define such numbers, namely by **defining an equivalence relation $\sim$ on** $\mathbb{N}^2$, by

$$(m, n) \sim (k, l) \overset{\text{def}}{\iff} m + l = k + n$$

That this is an equivalence relation is shown as follows: for all $(m, n), (k, l), (p, q) \in \mathbb{N}^2$ we have

(1) $m + n = m + n \implies (m, n) \sim (m, n)$ (reflexivity)

(2) $(m, n) \sim (k, l) \implies m + l = k + n \implies k + n = m + l \implies (k, l) \sim (m, n)$ (symmetry)

(3) $(m, n) \sim (k, l)$ and $(k, l) \sim (p, q) \implies m + l = k + n$ and $k + q = p + l \implies m + q + l = n + q + k = n + p + l \implies m + q = n + p \implies (m, n) \sim (p, q)$ (transitivity)

The equivalence class $[(m, n)]$ can then be used to **define an integer**. For example, we define -1 by

$$-1 \overset{\text{def}}{=} [(2, 3)]$$

The **set of all integers** $\mathbb{Z}$ is then naturally *defined as the set of all equivalence classes on* $\mathbb{N}^2$*, that is as the quotient set on* $\mathbb{N}^2$,

$$\mathbb{Z} \overset{\text{def}}{=} \mathbb{N}^2 / \sim$$

The arithmetic binary operations of **addition** $+$, **subtraction** $-$, and **multiplication** $\cdot$ on $\mathbb{Z}$ are defined as follows: for all $a = [(m, n)], b = [(k, l)] \in \mathbb{Z}$

$$\begin{aligned} a + b &= [(m, n)] + [(k, l)] = [(m + k, n + l)] \\ a - b &= [(m, n)] - [(k, l)] = [(m, n)] + [(l, k)] = [(m + l, n + k)] \\ ab &= [(m, n)][(k, l)] = [(mk + nl, ml + nk)] \end{aligned}$$

and a **partial order** $\leq$ given by

$$a = [(m,n)] \leq b = [(k,l)] \iff m + l \leq n + k$$

We will show in a theorem below that these operations are well defined (i.e. do not depend on the choice of representatives of the equivalence classes in $\mathbb{N}^2/\sim$). The integers are thus an algebraic structure $(\mathbb{Z}, \leq, +, -, \cdot)$ and a partially ordered set.

If we wish to have $\mathbb{N} \subseteq \mathbb{Z}$, then we will first need to embed the natural numbers as we have constructed them into $\mathbb{Z} = \mathbb{N}^2/\sim$ as we have constructed them. The canonical embedding $f : \mathbb{N} \hookrightarrow \mathbb{Z}$ is

$$f(n) = [(n,0)]$$

which is indeed an embedding: if $m, n \in \mathbb{N}$, then $f(m) = f(n)$ iff $[(m,0)] = [(n,0)]$ which implies $m = n$, since if $[(m,0)] = [(n,0)]$ then we clearly have $(n,0) \sim (m,0)$, or $m = m + 0 = n + 0 = n$. Moreover, f preserves the algebraic structure of $\mathbb{N}$:

$$f(m+n) = [(m+n,0)] = [(m,0)] + [(n,0)] = f(m) + f(n)$$

so that indeed $\mathbb{N} \overset{f}{\hookrightarrow} \mathbb{Z}$. When considering the subset $f(\mathbb{N})$ of $\mathbb{Z}$ we ususally write $\mathbb{N} \subseteq \mathbb{Z}$, strictly incorrectly of course, but in keeping with the intuitive notion of the natural numbers being a subset of the integers.

Proposition 4.1 *The binary operations of addition $+$, subtraction $-$, and multiplication $\cdot$ on $\mathbb{Z}$ are* **well defined**. *That is, if $[(m,n)] = [(m',n')]$ and $[(k,l)] = [(k',l')]$, then*

(1) $[(m+k, n+l)] = [(m'+k', n'+l')]$

(2) $[(m+l, n+k)] = [(m'+l', n'+k')]$

(3) $[(m,n)(k,l)] = [(m',n')(k',l')]$

(4) $[(m,n)] \leq [(k,l)] \iff [(m',n')] \leq [(k',l')]$

Proof: If $[(m,n)] = [(m',n')]$ and $[(k,l)] = [(k',l')]$, then $(m,n) \sim (m',n')$ and $(k,l) \sim (k',l')$, so that

$$m + n' = m' + n \tag{4.1}$$

and

$$k + l' = k' + l \tag{4.2}$$

(1) Adding these equations gives $m+k+n'+l' = m'+k'+n+l$, which shows that $(m+k, n+l) \sim (m'+k', n'+l')$, so that $[(m+k, n+l)] = [(m'+k', n'+l')]$, whence addition is well-defined. (2) That subtraction is well-defined follows from the fact that addition is. (3) We must show that $[(m,n)(k,l)] = [(m',n')(k',l')]$, i.e. that $[(mk+nl, ml+nk)] = [(m'k'+n'l', n'k'+m'l')]$, or equivalently $(mk+nl, ml+nk) \sim (m'k'+n'l', n'k'+m'l')$, or equivalently

$$mk+nl+n'k'+m'l' = m'k'+n'l'+ml+nk$$

From (4.1) it follows that $(m+n')(k+k') = (m'+n)(k+k')$ and $(m'+n)(l+l') = (m+n')(l+l')$, or

$$mk+n'k'+n'k+mk' = m'k'+nk+nk'+m'k \tag{4.3}$$
$$m'l'+nl+nl'+m'l = ml+n'l'+n'l+ml' \tag{4.4}$$

while from (4.2) it follows that $(m+m')(k+l') = (m+m')(k'+l)$ and $(n+n')(k'+l) = (n+n')(k+l')$, or

$$mk+m'l'+m'k+ml' = m'k'+ml+mk'+m'l \tag{4.5}$$
$$n'k'+nl+nk'+n'l = nk+n'l'+n'k+nl' \tag{4.6}$$

Adding (4.3) and (4.4) gives

$$\begin{aligned}(mk+nl+n'k'+m'l') + (n'k+mk'+nl'+m'l)\\ = (m'k'+n'l'+ml+nk) + (n'l+ml'+nk'+m'k)\end{aligned} \tag{4.7}$$

while adding (4.5) and (4.6) gives

$$\begin{aligned}(mk+nl+n'k'+m'l') + (n'l+ml'+nk'+m'k)\\ = (m'k'+n'l'+ml+nk) + (n'k+mk'+nl'+m'l)\end{aligned} \tag{4.8}$$

Letting $A = mk+nl+n'k'+m'l'$, $B = m'k'+n'l'+ml+nk$, $C = n'k+mk'+nl'+m'l$ and $D = n'l+ml'+nk'+m'k$, equations (4.7) and (4.8) can be rewritten

$$A+C = B+D \tag{4.9}$$
$$A+D = B+C \tag{4.10}$$

From these we get $(A+D)+D = (B+C)+D = (B+D)+C = (A+C)+C$, which implies $C+C = D+D$ by the cancellation law of addition for $\mathbb{N}$, Theorem 3.12. But then by the cancellation law of multiplication we have

$2C = C + C = D + D = 2D \implies C = D$. But if $C = D$, then (4.9) implies that $A + C = B + C$, whence, again by the cancellation law of addition, we have $A = B$, or

$$mk + nl + n'k' + m'l' = m'k' + n'l' + ml + nk$$

This means that $(mk+nl, ml+nk) \sim (m'k'+n'l', n'k'+m'l')$, or $(m,n)(k,l) \sim (m',n')(k',l')$, which means $[(m,n)(k,l)] = [(m',n')(k',l')]$. (4) Finally, if $[(m,n)] \leq [(k,l)]$, then by definition we have

$$m + l \leq n + k \tag{4.11}$$

and by the fact that $(m,n) \sim (m',n')$ and $(k,l) \sim (k',l')$ we have

$$\begin{aligned} m + n' &= m' + n & (4.12)\\ k + l' &= k' + l & (4.13) \end{aligned}$$

Adding (4.12) and (4.13) and using (4.11) gives

$$\begin{aligned} m' + l' + n + k &= m + l + n' + k' \\ &\leq n + k + n' + k' \end{aligned}$$

so by cancellation we have $m' + l' \leq n' + k'$, or $[(m',n')] \leq [(k',l')]$, and $\leq$ is well-defined. ■

Lemma 4.2 *If we define the binary operations $+$ and $\cdot$ on $\mathbb{N}^2$ by*

$$\begin{aligned} (a,b) + (c,d) &= (a + c, b + d) \\ (a,b)(c,d) &= (ac + bd, ad + bc) \end{aligned}$$

for all $(a,b),(c,d),(e,f) \in \mathbb{N}^2$, then these operations satisfy

1. $(a,b) + (c,d) = (c,d) + (a,b)$
2. $(a,b)(c,d) = (c,d)(a,b)$

   } (*commutativity*)
3. $((a,b) + (c,d)) + (e,f) = (a,b) + ((c,d) + (e,f))$
4. $((a,b)(c,d))(e,f) = (a,b)((c,d)(e,f))$

   } (*associativity*)
5. $\exists (0,0) \in \mathbb{N}^2$ *s.t.* $(a,b) + (0,0) = (0,0) + (a,b) = (a,b)$ (*additive identity*)
6. $\exists (1,0) \in \mathbb{N}^2$ *s.t.* $(a,b)(1,0) = (1,0)(a,b) = (a,b)$ (*multiplicative identity*)
7. $((a,b) + (c,d))(e,f) = (a,b)(e,f) + (c,d)(e,f)$ *and* $(a,b)((c,d) + (e,f)) = (a,b)(c,d) + (a,b)(e,f)$

   } (*distributivity*)

Proof: By the properties of addition and multiplication in $\mathbb{N}$ we have: **1.** $(a,b)+(c,d) = (a+c,b+d) = (c+a,d+b) = (c,d)+(a,b)$. **2.** $(a,b)(c,d) = (ac+bd,ad+bc) = (ca+db,da+cb) = (c,d)(a,b)$. **3.** $((a,b)+(c,d))+(e,f) = (a+c,b+d)+(e,f) = ((a+c)+e,(b+d)+f) = (a+(c+e),b+(d+f)) = (a,b)+(c+e,d+f) = (a,b)+((c,d)+(e,f))$. **4.** $((a,b)(c,d))(e,f) = (ac+bd,ad+bc)(e,f) = ((ac+bd)e+(ad+bc)f,(ad+bc)e+(ac+bd)f) = (a(ce+df)+b(de+cf),a(de+cf)+b(ce+df)) = (a,b)(ce+df,de+cf) = (a,b)((c,d)(e,f))$. **5.** For all $(a,b) \in \mathbb{N}_0^2$ we have $(a,b) = (a+0,b+0) = (a,b)+(0,0) = (0+a,0+b) = (0,0)+(a,b)$. **6.** For all $(a,b) \in \mathbb{N}_0^2$ we have $(a,b)(1,0) = (a1+b0,a0+b1) = (a,b)$ and $(1,0)(a,b) = (1a+0b,1b+0a) = (a,b)$. **7.** $((a,b)+(c,d))(e,f) = (a+c,b+d)(e,f) = ((a+c)e+(b+d)f,(a+c)f+(b+d)e) = (ae+ce+bf+df,af+cf+be+de) = (ae+bf,af+be)+(ce+df,cf+de) = (a,b)(e,f)+(c,d)(e,f)$. By commutativity and by the argument just made we also have $(a,b)((c,d)+(e,f)) = ((c,d)+(e,f))(a,b) = (c,d)(a,b)+(e,f)(a,b) = (a,b)(c,d)+(a,b)(e,f)$. ■

Theorem 4.3 (Arithmetic Properties of $\mathbb{Z}$) *For all $a,b,c \in \mathbb{Z}$ we have*

1.	$a+b \in \mathbb{Z}$	($\mathbb{Z}$ *is closed under* $+$, $-$ *and* $\cdot$)	($\mathbb{Z}$ *is a commutative ring*)
2.	$a-b \in \mathbb{Z}$		
3.	$ab \in \mathbb{Z}$		
4.	$a+b = b+a$	(*commutativity*)	
5.	$ab = ba$		
6.	$(a+b)+c = a+(b+c)$	(*associativity*)	
7.	$(ab)c = a(bc)$		
8.	$\exists 0 \in \mathbb{Z}$ *s.t.* $a+0 = 0+a = a$	(*additive identity*)	
9.	$\exists 1 \in \mathbb{Z}$ *s.t.* $a1 = 1a = a$	(*multiplicative identity*)	
10.	$\forall a \in \mathbb{Z}$, $\exists! b \in \mathbb{Z}$ *s.t.* $a+b = 0$	(*additive inverse*)	
11.	$(a+b)c = ac+bc$ *and* $a(b+c) = ab+ac$	(*distributivity*)	
12.	$(-1)a = -a$		

Proof: Properties **1-3** follow from the definitions of $+$, $-$ and $\cdot$ on $\mathbb{Z}$ and the fact that they are well-defined by P 4.1, while **4-9** and **11** follow from the lemma and the fact that $+$, $-$ and $\cdot$ are well-defined, since by P 4.1 we can restate each statement in the lemma in terms of equivalence classes $[(a,b)]$ instead of ordered pairs (a,b). For example, $0 := [(0,0)] \in \mathbb{Z}$ and

$1 := [(1,0)] \in \mathbb{Z}$. Finally, **10** follows from the fact that for all $a = [(m,n)] \in \mathbb{Z}$, there exists $b = [(n,m)] \in \mathbb{Z}$ such that

$$a+b = [(m,n)]+[(n,m)] = [(m+n, n+m)] = [(m+n, m+n)] = [(0,0)] = 0$$

since $(m+n, m+n) \sim (0,0)$ because $m+n+0 = 0+m+n$. Moreover, b is unique, for if there were any other $c \in \mathbb{Z}$ such that $a+b = a+c = 0$, then we would have

$$b = b+0 = b+(a+c) = (b+a)+c = 0+c = c$$

This b, the unique additive inverse of a, is usually denoted $-a$ and is called the **negative** of a. From this we may conclude that $\{1,2,\dots\} = f(\mathbb{N})$ and $\{-1,-2,\dots\} = \{-f(1'), -f(2'), \dots\} = -f(\mathbb{N})$ are negatives of each other, where f is the embedding of $\mathbb{N}_0 = \{0', 1', 2', \dots\}$ into $\mathbb{Z}$, and $\mathbb{Z} = f(\mathbb{N}_0) \sqcup -f(\mathbb{N})$, so that the integers extend the natural numbers by adding the negatives of all the nonzero naturals. Finally, **12.** $(-1)a = [(0,1)][(m,n)] = [(0m+1n, 0n+1m)] = [(n,m)] = -a$. ■

Definition 4.4 *If we take properties (1)-(3) as* definitions *of the binary operations of addition and multiplication,*

$$+ : \mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}, \qquad (a,b) \mapsto a+b \qquad (\textbf{addition})$$
$$\cdot : \mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}, \qquad (a,b) \mapsto ab \qquad (\textbf{multiplication})$$

along with the unary operation of negation

$$- : \mathbb{Z} \to \mathbb{Z}, \qquad a \mapsto -a \qquad (\textbf{negation})$$

then the other properties, (4)-(12), become the **axioms of an abstract ring***: if we remove the particular construction of $\mathbb{Z}$ out of equivalence classes of $\mathbb{N}^2$ given above and ask only for a set R to be equipped with $+$, $\cdot$, and $-$, and to satisfy the axioms, then we have the* definition of an algebraic ring. *Actually, to be perfectly correct, properties (4)-(12) define a* **commutative ring***, and a* **ring** *by itself only if we remove axiom (5). This allows, for example, square matrices to be considered a* noncommutative *ring.*

The next idea then is to situate $\mathbb{Z}$ within the class of all rings. MacLane himself co-invented the way to do this: make rings into a category, and then see how $\mathbb{Z}$ fits in that category. The answer is: $\mathbb{Z}$ is an 'initial element' in the category of rings, because of a certain 'universal property' it satisfies. We have thus arrived at the culmination of Aristotle's original idea of conceptualizing Plato's form of 'integer number,' namely as the ring of integers, an initial element in its category. ■

5 Rational Numbers

We now construct the rational numbers $\mathbb{Q}$ out of the integers $\mathbb{Z}$, as a quotient set of the cartesian product of the integers. The issue is similar to that with subtraction in constructing $\mathbb{Z}$ out of $\mathbb{N}$. Just as

$$2 - 3 = 3 - 4 = \cdots = -1$$

has many representations, all *equivalent* in the precise sense that, say,

$$\begin{aligned} (2,3) \sim (3,4) &\overset{\text{def}}{\iff} 2 + 4 = 3 + 3 \\ &\iff -1 \overset{\text{def}}{=} [(2,3)] \end{aligned}$$

just so we have

$$\frac{4}{6} = \frac{-8}{-12} = \cdots = \frac{2}{3}$$

and therefore we *define*

$$\begin{aligned} (4,6) \sim (-8,-12) &\overset{\text{def}}{\iff} 4 \cdot (-12) = 6 \cdot (-8) \\ &\iff \frac{2}{3} \overset{\text{def}}{=} [(4,6)] \end{aligned}$$

Definition 5.1 *Formally, define the relation $\sim$ on $\mathbb{Z}^2$ by*

$$(a,b) \sim (c,d) \iff (ad = bc \text{ and } b \neq 0,\ d \neq 0) \text{ or } (b = d = 0) \qquad \blacksquare$$

Proposition 5.2 *$\sim$ is an equivalence relation on $\mathbb{Z}^2$.*

Proof: For all $(a,b) \in \mathbb{Z}^2$, if $b = 0$ then $(a,b) \sim (a,b)$, while if $b \neq 0$, then $ab = ba$, so $(a,b) \sim (a,b)$. Moreover, if $(a,b) \sim (c,d)$, then if $b = d = 0$ we have $d = b = 0$ and so $(c,d) \sim (a,b)$, while if $b \neq 0$ and $d \neq 0$, then $ad = bc$, whence $cb = da$, so that $(c,d) \sim (a,b)$. Finally, if $(a,b) \sim (c,d)$ and $(c,d) \sim (e,f)$, and $b = d = f = 0$, then clearly $(a,b) \sim (e,f)$, while if $b \neq 0$, $d \neq 0$ and $f \neq 0$, then $ad = bc$ and $cf = ed \implies adf = bcf = bed \implies af = be \implies (a,b) \sim (e,f)$. $\blacksquare$

Definition 5.3 *Now, we can define the* **rational numbers** $\mathbb{Q}$ *as the quotient set of* $\mathbb{Z}^2$ *by* $\sim$*,*

$$\begin{aligned}\mathbb{Q} &= \mathbb{Z}^2/\sim \\ &= \{[(a,b)] \mid a,b \in \mathbb{Z} \text{ and } b \neq 0\}\end{aligned}$$

Moreover, $\mathbb{Q}$ *is endowed with the arithmetic operations of* **addition** $+$ *and* **multiplication** $\cdot$ *given by*

$$\begin{aligned}[(a,b)] + [(c,d)] &= [(ad+bc, bd)] \\ [(a,b)][(c,d)] &= [(ac,bd)]\end{aligned}$$

and a **partial order** $\leq$ *given by*

$$[(a,b)] \leq [(c,d)] \quad \overset{\text{def}}{\iff} \quad b,d \geq 0 \text{ and } ad \leq bc$$

which are also well-defined, as will be shown below. Moreover, $\mathbb{Q}$ *contains an* **additive identity** $0 = [(0,1')]$*, where* $1'$ *is the multiplicative identity of* $\mathbb{Z}$*, since for all* $p = [(a,b)] \in \mathbb{Q}$ *we have*

$$\begin{aligned}p + 0 = 0 + p &= [(a,b)] + [(0,1')] = [(a1 + b0, b1')] = [(a,b)] = p \\ &= [(a,b)] = [(0b + 1'a, 1'b)] = [(0,1')] + [(a,b)] = 0 + p\end{aligned}$$

Also, $\mathbb{Q}$ *contains a* **multiplicative identity***,*

$$1 \overset{\text{def}}{=} [(1',1')]$$

where $1'$ *is the multiplicative identity of* $\mathbb{Z}$*, since for any* $p = [(a,b)] \in \mathbb{Q}$ *we have*

$$p1 = [(a,b)][(1',1')] = [(a1', b1')] = [(a,b)] = [(1'a, 1'b)] = [(1',1')][(a,b)] = 1p$$

Finally, we endow $\mathbb{Q}$ *with the unary* **negative** *function* $-$ *and the unary* **multiplicative inverse** *function* $^{-1}$*, which send each* $[(a,b)] \in \mathbb{Q}$ *into*

$$-[(a,b)] \overset{\text{def}}{=} [(-a,b)]$$

and

$$[(a,b)]^{-1} \overset{\text{def}}{=} [(b,a)]$$

in $\mathbb{Q}$*, respectively, the last only if* $[(a,b)] \neq 0$*. Another way to say this is that each* $[(a,b)] \in \mathbb{Q}$ *has an additive inverse and each* $[(a,b)] \in F\backslash\{0\}$ *has a multiplicative inverse, as give above. Thus,* $\mathbb{Q}$ ***is an albegraic structure***

$$(\mathbb{Q}, \leq, -, ^{-1}, +, \cdot, 0, 1)$$

It remains to show that is is a field (every nonzero rational number has a multiplicative inverse), and an ordered set $(\mathbb{Q}, \leq)$. To show this, however, we first need to show that our addition and multiplication, as well as our order relation, are all well defined.

Remark 5.4 *We may* **embed the integers into the rationals** *in a way similar to the embedding of the naturals into the integers. The function $f : \mathbb{Z} \to \mathbb{Q}$, i.e. $f : \mathbb{Z} \to \mathbb{Z}^2/\sim$, given by $f(a) = [(a, 1)]$, is injective, since $f(a) = f(b)$ implies $[(a, 1)] = [(b, 1)]$, so $(a, 1) \sim (b, 1)$, or $a = a1 = b1 = b$. Thus, $\mathbb{Z} \overset{f}{\hookrightarrow} \mathbb{Q}$, though we usually write $\mathbb{Z} \subseteq \mathbb{Q}$.* ■

Proposition 5.5 *The unary negative $-$ and multiplicative inverse $^{-1}$ functions, the binary addition $+$ and multiplication $\cdot$ functions, and the partial order relation $\leq$ on $\mathbb{Z}^2/\sim$ are all well defined. That is, if $[(a, b)] = [(a', b')]$ and $[(c, d)] = [(c', d')]$ in $\mathbb{Z}^2/\sim$, then*

1. $-[(a, b)] = -[(a', b')]$
2. $[(a, b)]^{-1} = [(a', b')]^{-1}$
3. $[(ad + bc, bd)] = [(a'd' + b'c', b'd')]$
4. $[(ac, bd)] = [(a'c', b'd')]$
5. $[(a, b)] \leq [(c, d)] \iff [(a', b')] \leq [(c', d')]$

Proof: The first two follow easily from the definitions. If $[(a, b)] = [(a', b')]$ and $[(c, d)] = [(c', d')]$ are rational numbers, then $b, b'd, d' \neq 0$ and $(a, b) \sim (a', b')$ and $(c, d) \sim (c', d')$, so $ab' = a'b$ and $cd' = c'd$. **1.** Consequently, $-ab' = -a'b$, so $(-a, b) \sim (-a', b')$, and therefore $-[(a, b)] = [(-a, b)] = [(-a', b')] = -[(a', b')]$. **2.** Similarly, if $[(a, b)] \neq 0 = [(0, 1)]$, we have $a = a1 \neq b0 = 0$, so $[(b, a)]$ is defined, and likewise $[(b', a')]$ is defined. But if $ab' = a'b$, then certainly $b'a = ba'$, whence $(b, a) \sim (b', a')$, and so $[(a, b)]^{-1} = [(b, a)] = [(b', a')] = [(a', b')]^{-1}$. **3.** Note that

$$acb'd' = a'c'bd \tag{5.1}$$

whence, first of all, $(ac, bd) \sim (a'c', b'd')$, which means

$$[(ac, bd)] = [(a'c', b'd')]$$

whence multiplication is well defined. But (5.1) also implies

$$acb'd' + bdb'd' = a'c'bd + bdb'd'$$

or $(ac+bd)b'd' = (a'c'+b'd')bd$, which is the definition of

$$(ac+bd, bd) \sim (a'c'+b'd', b'd')$$

so that

$$[(ad+bc, bd)] = [(a'd'+b'c', b'd')]$$

whence addition is well-defined. Moreover, if $[(a,b)] \leq [(c,d)]$, then $b, d \geq 0$, in fact $b, d > 0$ since $b, b', d, d' \neq 0$, and

$$ad \leq bc \tag{5.2}$$

We must show that $a'd' \leq b'c'$ and $c', d' \geq 0$. But since

$$ab' = a'b \tag{5.3}$$
$$cd' = c'd \tag{5.4}$$

we have

$$a'd'bd \overset{(5.3)}{=} ad'b'd \overset{(5.2)}{\leq} bd'b'c \overset{(5.4)}{=} bdb'c'$$

By cancellation, therefore, since $b, d > 0$, we have $a'd' \leq b'c'$. It remains to show that $b', d' \geq 0$. Now, note that for all $(a,b) \in \mathbb{Z}^2$ with $b > 0$ we have $(a,b) \sim (-a,-b)$, so if we choose $(-a,-b)$ as a representative of $[(a,b)]$, then clearly for any other $(c',d') \in [(c,d)]$ we'll have $-ad' \leq -bc'$, and yet $-b < 0$. Thus, we must additionally require that the representatives (a',b') and (c',d') satisfy $b', d' > 0$ from the beginning. Under this assumption the theorem holds. ■

Theorem 5.6 (Arithmetic Properties of $\mathbb{Q}$) *For all $a, b, c \in \mathbb{Q}$ we have*

1.	$a + b \in \mathbb{Q}$	($\mathbb{Q}$ *is closed under* $+$, $-$ *and* $\cdot$)	($\mathbb{Q}$ *is a field*)
2.	$a - b \in \mathbb{Q}$		
3.	$ab \in \mathbb{Q}$		
4.	$a + b = b + a$	(*commutativity*)	
5.	$ab = ba$		
6.	$(a + b) + c = a + (b + c)$	(*associativity*)	
7.	$(ab)c = a(bc)$		
8.	$\exists 0 \in \mathbb{Q}$ *s.t.* $a + 0 = 0 + a = a$	(*additive identity*)	
9.	$\exists 1 \in \mathbb{Q}$ *s.t.* $a1 = 1a = a$	(*multiplicative identity*)	
10.	$\forall a \in \mathbb{Q}$, $\exists! b \in \mathbb{Q}$ *s.t.* $a + b = 0$	(*additive inverse*)	
11.	$\forall a \in \mathbb{Q} \backslash \{0\}$, $\exists! b \in \mathbb{Q}$ *s.t.* $ab = 1$	(*multiplicative inverse*)	
12.	$(a + b)c = ac + bc$ *and* $a(b + c) = ab + ac$	(*distributivity*)	
13.	$(-1)a = a(-1) = -a$		
14.	$\forall a, b, \in \mathbb{Q}$, $a \leq b$ *or* $b \leq a$	($\leq$ *is total order*)	

Thus, $\mathbb{Q}$ is a structure $(\mathbb{Q}, \leq, -, ^{-1}, +, \cdot, 0, 1)$ which is a totally ordered field.

Proof: Properties **1** and **3** follow from the definitions of $+$ and $\cdot$ on $\mathbb{Q}$ and the fact that they are well-defined by Proposition 5.5, while **2** follows from **10** and **1**, so that we can say $a - b = a + (-b)$, and $(-b)$ is a unique rational. **4**-**7** and **12** follow from the properties of $\mathbb{Z}$, Theorem 4.2, which translate into addition and multiplication on $\mathbb{Z}^2$, analogously to Lemma 4.2, and the fact that the operations on equivalence classes in $\mathbb{Z}^2$ are well-defined by Proposition 5.5. The existence of $0 = [(0', 1')]$ and $1 = [(1', 1')]$, where $0', 1' \in \mathbb{Z}$ we described in the introduction, so all that remains is the existence and uniqueness of the additive and multiplicative inverses. Towards this end, let $p = [(a, b)] \in \mathbb{Q}$, and note that $q = [(-a, b)]$ satisfies

$$p + q = [(a, b)] + [(-a, b)] = [(ab - ab, bb)] = [(0, bb)] = [(0, 1')] = 0$$

Moreover, q is unique, for if there were any other $r \in \mathbb{Q}$ with $p + r = 0$ we would have

$$q = q + 0 = q + (p + r) = (q + p) + r = 0 + r = r$$

Now, if $p \neq 0$, which is equivalent to $a, b \neq 0$, then the rational number $q = [(b, a)]$ satisfies

$$pq = [(a, b)][(b, a)] = [(ab, ba)] = [(ab, ab)] = [(1', 1')] = 1$$

because $(ab, ab) \sim (1', 1')$ because $ab = ab1'$. Moreover, q is unique, since if r is any other rational with $pr = 1$, then

$$q = q1 = q(pr) = (qp)r = 1r = r$$

13. Finally, for all $p = [(a, b)] \in \mathbb{Q}$ we have

$$(-1)p = [(-1', 1')][(a, b)] = [((-1)'a, 1b)] = [(-a, b)] = -p$$

and

$$p(-1) = [(a, b)][(-1', 1')] = [(a(-1'), 1b)] = [(-a, b)] = -p$$

as well. This completes the proof.

14. Since $\mathbb{Z}$ is totally ordered, for all $p = [(a, b)], q = [(c, d)] \in \mathbb{Q}$ we have $ac \leq bd$ or $bd \leq ac$, whence either $p \leq q$ or $q \leq p$, which shows that $\mathbb{Q}$ is totally ordered. ■

Definition 5.7 *If we take properties (1)-(3) as* definitions *of the binary operations of addition and multiplication,*

$$+ : \mathbb{Q} \times \mathbb{Q} \to \mathbb{Q}, \qquad (a, b) \mapsto a + b \qquad (\textbf{addition})$$
$$\cdot : \mathbb{Q} \times \mathbb{Q} \to \mathbb{Q}, \qquad (a, b) \mapsto ab \qquad (\textbf{multiplication})$$

along with the unary operation of negation

$$- : \mathbb{Q} \to \mathbb{Q}, \qquad a \mapsto -a \qquad (\textbf{negation})$$

then the other properties, (4)-(13), become the **axioms of an abstract field***: if we remove the particular construction of $\mathbb{Q}$ out of equivalence classes of $\mathbb{Z}^2$ given above and ask only for a set F to be equipped with $+$, $\cdot$, and $-$, and to satisfy the axioms, then we have the* definition of an algebraic field*.*

The next question concerning fields is not categorical, as with the integers. The next question concerns ideas from field theory itself, namely various extensions of the field $\mathbb{Q}$, which is typically studied in the second semester of an abstract algebra course. The most important extension of $\mathbb{Q}$ for us is $\mathbb{R}$, the real number field. ■