Introduction to Set Theory

Sean Murphy

April 4, 2021

Contents

Chapter 1

The Language of Zermelo-Fraenkel Set Theory

We begin our investigation of sets with a brief recap of first order logic. While set theory tries to give a fundamental basis for mathematics, we need to use some notation from other areas of logic to be concrete in our definitions. This is where the language of mathematical logic comes in handy to us.

1.1 Logical Symbols

To construct a formal language such as mathematics, we need to define a formal alphabet comprised The following are all **logical symbols**:

$$=, x, y, (,), \wedge, \vee, \neg, \implies, \iff, \forall, \exists$$

The first one is our good friend the equals sign, where $x = 4$ means that the variable x is the same as the quantity 4. The parentheses parse the language and let us concatenate phrases together. Similarly, the $\wedge$ and $\vee$ signs connect the truth value of two formulas together, but either requiring both to be true or at least one to be true, respectively. The $\neg$ switches the truth value of a phrase, and the $\implies$ and $\iff$ symbols connect to phrases together, where $A \implies B$ means A is true only if B is true and $A \iff B$ means A is true if and only if B is true. The last two are actually existential quantifiers, where $\forall$ means 'for every' or 'for all' and $\exists$ means 'there exists' or 'there is.' We will see the usefulness of this in a bit.

1.2 Vocabulary

The **vocabulary** of mathematics is comprised of many symbols:

$$+, -, \times, \div, \in, \subseteq, \leq \div, \text{etc.}$$

One of the beauties of set theory is that everything is a set. In particular, we can theoretically express all the above operations in terms of just these symbols. Of course, $x \in A$ means that x is a member of A and $A \subseteq B$ means A is a subset of B. That is, $x \in A$ implies that $x \in B$. It is possible there is a $y \in B$ where $y \notin A$ (In fact, $x \subseteq a$ can be replaced by $\forall t(t \in x \implies t \in a.)$

1.3 Formulas

With these notions of connectives, quantifiers, and variables, we can talk about general first order formulas. A **formula** ψ can be thought of as a sequence of symbols with a truth value associated to it. It's important that ψ is well-formed. For instance, $\psi : \exists x(x < 4)$ is a well-defined formula but $\varphi : \exists y \implies)($ is not. If ψ, φ are formulas in this language then so are $\neg\psi, \psi \wedge \varphi, \psi \vee \varphi, \psi \implies \varphi, \psi \iff \varphi$.

1.4 Examples

Let's now consider some statements in English that can be translated into this language. For example, *there is no largest natural number*. This means that for any natural number n, we can find another natural number m that is greater than n. Taking $\mathbb{N}$ to denote the natural numbers (which we will cover in detail later) we have the following equivalent mathematical statement:

$$\forall n \in \mathbb{N}(\exists m \in \mathbb{N}((m > n))$$

OK- that was not too hard. Let's tackle a more complex statement, and this time we will read in the first order formula first:

$$\forall x \in \mathbb{N}(\exists p \in \mathbb{N}(p > x \wedge \forall y \in \mathbb{N}(\forall z \in \mathbb{N}((y > 1 \wedge z > 1) \implies \neg(y \cdot z = p))))$$

Although it looks complicated, the meaning is rather intuitive. Given any natural number x, there is a natural number p such that the following holds: p is greater than x and for any two natural numbers y and z that are both greater than 1, it cannot be the case that y times

z is p. Okay, so that means p is prime. But the outer layer of this statements says that this holds for all naturals x, and so p can be arbitrarily large. That is to say, there are infinitely many primes, neat!

1.5 Truth versus Proof

While the above statements are **true**, in general a well-formed first-order formula need not be true. For instance, $\exists y \in \mathbb{R}(\forall x \in \mathbb{R}(x < y))$ is not true, since there is no largest real number.

In mathematics, many statements that we wish to **prove** are of the form $\psi \implies \varphi$. That is to say, under the *assumption* that ψ is true, then φ is also true. To *prove* such a mathematical statement is to make a sequence of valid assertions under the assumption that ψ is true, and deduce that φ holds. Another (equivalent) method is to use the method of contraposition: if we assume that φ is false, then it must be that ψ is false. The equivalence holds based on the truth values of the two formulas. In English, consider the following statement: *If it's sunny then the sky is blue.* This makes perfect sense to us, but it can be thought of as a formal statement. Consider the *contrapositive* of this statement: *If the sky is not blue then it's not sunny.* See the equivalence? If the sky is not blue, then it must be the case that it's not sunny (if it were, then the sky would be blue.) It is assumed that the reader has familiarity with proofs, but it never hurts to have a refresher. With the language of mathematics seen in the previous section, let us now talk about our object of focus, sets.

1.6 The Empty Set

In most college-level mathematics courses you often see references to sets, for instance the unit circle $\{(x, y)\colon x^2 + y^2 \leq 1\}$ or a finite collection $\{1, 2, 3\}$. Some courses may even give a brief introduction to what sets are. In general, a **set** is a collection of other objects. In set theory every object is itself a set, and so a set can be thought of as a collection of other sets. The sets x in a set A are called the **members** of A. This relationship can be denoted as $x \in A$. Note that x is itself a set, and so there may be members in x as well. The **order** in which the elements (members) of a set appear in the description of a set do not matter. This is different than *ordered-pairs-* which we will talk about later- which only only refer to the ordering of elements in the Cartesian Product of two sets.

The above definition may sound recursive and self-referenced. If a set is a collection of other sets, then how do we know when we've reached the "bottom level" of a set? We get by this fact with the first axiom of set theory.

Axiom 1 (Empty set Axiom)**.** *There exists a set having no members, and we denote it with the symbol $\varnothing$. In first-order logic, this is: $\exists A(\forall x(x \notin A))$.*

The empty set can be thought of as a box with nothing inside of it, or with curly braces $\{\}$. One may ask the question, is the empty set in the empty set? No, $\varnothing \notin \varnothing$, since $\varnothing$ has no members. The set $\{\varnothing\}$ is $\{\{\}\}$ whereas the empty set is $\{\}$. Recall from the section on logic that well-formed formulas have a truth value. In math, the axioms are *assumed* to be true. These are statements that we take for granted, the ones we shall use to prove everything else.

1.7 Power Sets and Equality

Well, okay, let's assume that we have an empty set. Can we do anything with this? I mean, without any other true statements, all we know is that there is a set with no members. How is this useful? Consider the next axiom of Zermelo-Fraenkel set theory:

Axiom 2 (Power Set Axiom)**.** *For any set A, there is a set whose members are exactly the subsets of A. In first-order logic, this is: $\forall A(\exists B(\forall x(x \in B \implies x \subseteq A)))$. We denote the power set of A as $\mathcal{P}(A)$ or simply $\mathcal{P}A$.*

What is the power set of the empty set? It is the collection of sets x such that $x \subseteq \varnothing$. Well, the only subset of the empty-set is itself, since any non-empty set has a member y that lies outside of $\varnothing$. In conclusion, $\mathcal{P}(\varnothing) = \{\varnothing\}$. Of course, we can repeat this process: the power set of $\{\varnothing\}$ is $\{\varnothing, \{\varnothing\}\}$. The empty set is a subset of every set, and $\{\varnothing\}$ is a subset of itself. In fact, every set is a subset of itself.

While we can continue taking power sets of this fashion, it would be nice if we could work more closely with two arbitrary sets.

Axiom 3 (Extensionality Axiom)**.** *If two sets have exactly the same elements (members) then they are the same set. That is, $\forall A, B(\forall x(x \in A \iff x \in B) \implies A = B)$.*

This is an axiom that we take for granted, just like the average gradeschooler takes for granted that $2 = 2$. Since the order of the elements doesn't matter, the sets $\{1, 2, 3\}$ and $\{2, 3, 1\}$ are equal. The set $\{1, 2\}$ is different. By extentionality, to *prove* two sets A and B are equal, it suffices to show that $A \subseteq B$ and $B \subseteq A$.

1.8 Pairings and Unions

What if we have two sets, and want to consider the set with just those two sets as members? For instance, we know that $\varnothing$ and $\{\varnothing, \{\varnothing\}\}$ both exists. The power set axiom only let's us generate more power sets, so the next set of this schema is of four elements. How do we get just three elements?

Axiom 4 (Pairing Axiom)**.** *For any two sets a and b, there is a set C having a and b as elements. That is to say, $\forall a, b(\exists C(x \in C \implies x = a \vee x = b))$.*

Thus, given $\varnothing$ and $\{\varnothing, \{\varnothing\}\}$, we can create the set $\{\varnothing, \{\varnothing, \{\varnothing\}\}\}$ by pairing. Furthermore, we can pair with $\{\varnothing\}$ to get $\{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\}$, a set of three elements- not bad. We might also like to obtain a set of elements from a collection of other sets. We denote the **union** of a and b as $a \cup b$. Given the sets $\{3, 5, 6\}$ and $\{1, 2, 4\}$, the union is $\{3, 5, 6\} \cup \{1, 2, 4\} = \{1, 2, 3, 4, 5, 6\}$. We use the following axiom.

Axiom 5 (Union Axiom)**.** *For any set A there is a set whose members are those sets belonging to the sets inside A. That is, $\forall A(\exists B(\forall x(x \in B \iff (\exists a \in A(a \in a)))))$. We denote such such set B as $\bigcup A$.*

The union axiom says that $x \cup y$ exists for any sets x and y. Using the pairing axiom with A and B we obtain $\{A, B\}$. Applying the union axiom here gives $\bigcup\{A, B\} = A \cup B$. For if $x \in \bigcup\{A, B\}$ then $x \in z \in A$ or $x \in z \in B$ for some z. In either case $x \in A \cup B$. Likewise $x \in A \cup B$ means that $x \in \bigcup\{A, B\}$ since x is a member of some member.

1.9 Subsets and Schema of Separation

We haven't yet clarified whether subsets exists. While the pairings and unions let us "build" up sets, we'd also like a notion of obtaining "smaller" sets. For instance, we can calculate the union of a set, $\bigcup A$. How do we calculate the intersection of a set, $\bigcap A$?

Axiom 6 (Subset Axiom)**.** *For any set a and any formula $\psi(x)$ (that is, ϕ depends on x) there is a set b composed of all elements of a that satisfy ψ, written $b \subseteq a$. In first-order logic that is, $\forall a(\exists b(x \in b \iff \psi(x) \wedge x \in a))$.*

Think of ψ as a function where $\psi(x)$ is true depending on what x is. To get a subset of A based on $\psi(x)$ is to consider the set $\{x \colon x \in A \wedge \psi(x)\}$. For example, consider the set

$\{5, 121, 1729\}$. Combined with the formula $\psi(x) : x \neq 5$ we obtain the subset $\{121, 1729\}$. What about the even numbers? This is the set $2\mathbb{Z} \subseteq \mathbb{Z}$ where $x \in 2\mathbb{Z} \iff z \in \mathbb{Z} \land \exists n \in \mathbb{Z}(2n = x)$.

The subset axiom tells us that the intersection of two sets exist. Let A and B be two sets. The **intersection** of A and B is the set of all elements that lie in A and in B, and is denoted as $A \cap B$. That is, $\forall x((x \in A \land x \in B) \iff (x \in A \cap B))$. For a single set A, we define the intersection of $\bigcap A$ as the set $\{x \colon \forall a \in A(x \in a)\}$ by the subset axiom. As a remark, in other texts (and online) this axiom may referred to as the *Axiom Schema of Separation*. In principle, they mean the same thing.

1.10 Recap

We've seen a few axioms (namely, six out of nine total) and we've gotten some easy definitions out of the way. It's good to take a step back and reflect on what we've done so far. With the definition of the set, we have constructed some axioms about sets that we *assume* to be true. We will see in the following chapters how these axioms help us to create other mathematical objects that we already know of, and these axioms' limitations. With some concrete examples, we can see how all of mathematics can be interpreted in the language of set theory.

Chapter 2

Sets as Mathematical Objects

So all mathematical objects can be represented as sets- is this really true? I mean, surely a function is innately different than a set? What about a group, ring or field? Let's examine some familiar terminologies.

2.1 Ordered Pairs

If you've even done first year algebra, you've seen the graphs of various functions and examined points like $(1,2)$. This is the point that lies 1 unit in the x-direction and 2 units in the y-direction. It is different than the point $(2,1)$, since this lies 2 units along the x-direction and 1 unit in the y-direction. The order of the numbers matters. Now this may sound contradictory to what was said earlier, that the order of the members of a set don't matter. Well, let's be a little careful about what we're talking about.

Definition 1 (Ordered Pair)**.** *An ordered pair* $\langle a,b\rangle$ *is the set* $\{\{a\},\{a,b\}\}$.

Nice, so $\langle 1,2\rangle = \{\{1\},\{1,2\}\} \neq \{\{2\},\{2,1\}\} = \langle 2,1\rangle$, which is what we set out to show earlier, where $\langle 1,2\rangle = (1,2)$ is the proper notation. So why is it called an ordered-pair?

Theorem 1 (Ordered-Pairs are Well-Defined)**.** $\{\{x\},\{x,y\}\} = \{\{a\},\{a,b\}\}$ *if and only if* $x = a$ *and* $y = b$.

Proof. If $x = a$ and $y = b$ the result is true by extensionality. On the other hand suppose that $c = \{\{x\},\{x,y\}\} = \{\{a\},\{a,b\}\} = d$. Assume for now that $x \neq y$ and $a \neq b$. If

$\{x\} \in C = D$ then $\{x\} = \{a\}$ since they are each one element sets and $\{x, y\} \neq \{x\}$ since $x \neq y$ (that is, $y \in \{x, y\}$ but $y \notin \{x\}$. Apply extensionality.) This leads to $\{x, y\} = \{a, b\}$. By extensionality, these are the same elements. But we've already assumed $x = a$. Thus $\{a, y\} = \{a, b\}$. Again, $a \neq b$ and so $y = b$. If $x = y$ and $a = b$ then $\{a, b\} = \{a, a\} = \{aa\} = \{a\}$. This means $d = \{\{a\}\} = c$. But c is $\{\{x\}, \{x, y\}\}$. Since c has one member, $x = y$. Then $\{\{a\}\} = \{\{x\}\} \implies \{a\} = \{x\} \implies a = x$. □

The previous theorem/proof may seem pedantic, but the lesson is entirely worth it: we've carefully applied the axioms to arrive at the truth of the theorem via a sequence of logical conclusions. In lieu of Theorem 1, we see that ordered-pairs are well-defined and there is no ambiguity to this definition.

2.2 Cartesian Products

If you know linear algebra, you know that $\mathbb{R}^2$ can be viewed as a vector space with vectors of the form (x_1, x_2) where x_1, x_2 are real numbers. Like before, the vector $(1, 0)$ is different from $(0, 1)$, since these are the standard basis vectors for this vector. What is $\mathbb{R}^2$ as a set then? Consider the following definition.

Definition 2 (Cartesian product)**.** *If A and B are sets the Cartesian product of A and B, written $A \times B$, is the set $\{\langle x, y\rangle : x \in A, y \in B\}$. Formally, C is given by the following*

$$\forall A, B(\exists C((x \in A \wedge y \in B) \iff \langle x, y\rangle \in C))$$

The vector space $\mathbb{R}^2$ is shorthand for $\mathbb{R} \times \mathbb{R}$, the Cartesian product of the real numbers with itself. Any vector $v \in \mathbb{R}^2$ is an ordered pair with coordinates specifying the component of each basis vector along v. The product $\mathbb{N} \times \{0, 1\}$ is the set of all ordered pairs $\langle n, i\rangle$ where n is a natural number and i is 0 or 1. That is,

$$\mathbb{N} \times \{0, 1\} = \{\langle 0, 0\rangle, \langle 0, 1\rangle, \langle 1, 0\rangle, \langle 1, 1\rangle, \langle 2, 0\rangle, \ldots, \langle m, 0\rangle, \langle m, 1\rangle, \langle m + 1, 0\rangle, \ldots\}$$

Lemma 1. *If $x, y \in C$ then $\langle x, y\rangle \in \mathcal{P}(\mathcal{P}(C))$, the power set of the power set of C.*

Proof. Suppose that x and y are members of C. Then $\{x, y\}$ is a set by the pairing axiom and $\{x, y\}$ is a subset of C (by the condition $z \in \{x, y\}$.) Since $\mathcal{P}(C)$ contains all subsets of C, it follows that $\{x, y\}$ is a member of $\mathcal{P}(C)$. Likewise, $\{x\}$ is a subset of C and thus a member of $\mathcal{P}(C)$. Finally, we can pair these sets to get $\{\{x\}, \{x, y\}\}$ which must be a subset of $\mathcal{P}(C)$, and thus a member of $\mathcal{P}(\mathcal{P}(C))$. Of course, this is just the ordered pair $\langle x, y\rangle$, so our hypothesis is true. □

With this is mind, we see that the Cartesian product of A and B can be expressed as the following:

$$A \times B = \{z \in \mathcal{P}(\mathcal{P}(C)) \colon \exists x, y(x \in A \wedge y \in B \wedge z = \{\{x\}, \{x, y\}\}\}$$

2.3 Relations

Your *relatives* are people who share a common family member with you. You say you are *related* to your siblings because you have the same parents. Much like the real world, mathematical relations can represent the similar concepts.

Definition 3 (Relation)**.** *For any two sets A and B, a subset $R \subseteq A \times B$ is called a relation from A to B.*

A relation is a set of ordered pairs. Given two sets A and B, we can define Div to be the divisibility relation: $\langle a, b\rangle \in \mathrm{Div} \subseteq A \times B \iff a$ divides b. For instance, if $A = \{2, 4, 5\}$ and $B = \{5, 10, 40\}$ then $\mathrm{Div} = \{\langle 2, 10\rangle, \langle 2, 40\rangle, \langle 4, 40\rangle, \langle 5, 5\rangle, \langle 5, 10\rangle, \langle 5, 40\rangle\}$. Whenever R is a relation and $\langle a, b\rangle \in R$ we may write aRb, which says that a is *related* to b via R.

Definition 4 (Domain, Range, Field)**.** *Let R be a relation. We write:*

(a) $\operatorname{dom} R = \{a \colon \exists b(\langle a, b\rangle \in R)\}$

(b) $\operatorname{ran} R = \{b \colon \exists a(\langle a, b\rangle \in R)\}$

(c) $\operatorname{fld} R = \operatorname{dom} R \cup \operatorname{ran} R$

The terms **domain** and **range** may sound familiar. If $R \subseteq \{0, 1\} \times \{0, 1, 2\}$ is the set $\{\langle 0, 1\rangle, \langle 1, 2\rangle\}$, then $\operatorname{dom} R = \{0, 1\}$ and $\operatorname{ran} R = \{1, 2\}$. Note that the 1 appearing in the domain of R comes from $\{0, 1\}$ whereas the 1 in range of R comes from the set $\{0, 1, 2\}$. These sets exist by the subset axiom.

Lemma 2. *If R is a relation, then $\operatorname{fld} R = \bigcup\bigcup R$.*

Proof. Suppose that x is a member of $\operatorname{fld} R$. It must be that x is in $\operatorname{dom} R$ or x is in $\operatorname{ran} R$, by definition. Without loss of generality suppose $x \in \operatorname{dom} R$ (the other case is analogous.) There exists y such that $\langle x, y\rangle = \{\{x\}, \{x, y\}\} \in R$. Taking the union of R obtains members of the members of R, so $\{x\} \in \bigcup R$. Taking another union, we see that $x \in \bigcup\bigcup R$. As x

was arbitrary, we see that $\operatorname{fld} R \subseteq \bigcup\bigcup R$. On the other hand, if x is a member of $\bigcup\bigcup R$ then there exists another set z such that $x \in z \in \bigcup R$. Indeed, this means there is a w such that $z \in w \in R$ as well. As R is a relation, w is of the form $\{\{a\}, \{a, b\}\}$, and so z is of the form $\{a\}$ or $\{a, b\}$. Thus, $x \in z$ means that x is either a or b, and in either case x lies in the field of R (in the case that $a = b$ above wee see that $x = a$ and $\langle x, x\rangle \in R$. In conclusion, $\bigcup\bigcup R \subseteq \operatorname{fld} R$ and combined with the previous result we have set equality. □

2.4 Functions

The function $f(x) = x^2$ takes in values x and outputs the square of x. It is a **function** because it has one output for every input. The inverse $x \mapsto \pm\sqrt{x}$ is not a function because there is ambiguity as to whether x gets mapped to $\sqrt{x}$ or $-\sqrt{x}$.

Definition 5 (Function). *A function $f\colon A \to B$ is a relation $f \subseteq A \times B$ such that* $\operatorname{dom} f = A$ *and for all $x \in \operatorname{dom} f$ there exists a unique $y \in \operatorname{ran} f$ such that $\langle x, y\rangle \in f$. We say that f maps from A to B. $\forall x \in A(\exists y \in B(\langle x, y\rangle \in f \wedge \forall z(\langle x, z\rangle \in f \implies z = y)))$.*

It is important in our definition that we ensure $\operatorname{dom} f = A$ so that every value in the domain has an output from f. Our square root function above is a **map** $f \subseteq \mathbb{R} \to \mathbb{R}$ where $\langle x, x^2\rangle \in f$ for any $x \in \mathbb{R}$. Apart from $x = 0$ or $x = 1$ we never see $\langle x, x\rangle \in f$. The connotations of domain, range, and field from relations carry over to functions as well, since all functions are also relations.

Definition 6. *For a set A and relations f and g, we write:*

(a) $f^{-1} \coloneqq \{\langle y, x\rangle :\ \langle x, y\rangle \in f\}$

(b) $f \circ g \coloneqq \{\langle x, z\rangle : \exists y(\langle x, y\rangle \in f \wedge \langle y, z\rangle \in g)\}$

(c) $f|_A \coloneqq \{\langle x, y\rangle :\ \langle x, y\rangle \in f \wedge x \in A\}$

(d) $f[\![A]\!] \coloneqq \{y\colon \exists x(x \in A \wedge \langle x, y\rangle \in f)\}$

You can think of f^{-1} as the *inverse* of f (although f may not be invertible.) $f \circ g$ is called the *composition* of f and g, ad $f|_A$ is the *restriction* of f to the set A: we only consider the pairs in f whose first component lies in A. We say that $f[\![A]\!]$ as the *image* of A, because it is precisely the set of elements that f maps A into. Likewise, we can think of $f^{-1}[\![A]\!]$ as the *preimage* of A under f. The set $s = \{\langle 1, 2\rangle, \langle 3, 1\rangle, \langle 1, 3\rangle, \langle 2, 2\rangle\}$ is not a function- $1s2$ and $1s3$, where s is viewed as a relation. Indeed, s is a relation as it is a subset of the Cartesian product $\{1,2\} \times \{1,2,3\}$. On the other hand, $r = \{\langle 1,3\rangle, \langle 3,2\rangle, \langle 2,1\rangle\}$ is a function. The inverse of s is $s^{-1} = \{\langle 2,1\rangle, \langle 1,3\rangle, \langle 3,1\rangle, \langle 2,2\rangle\}$. Taking t to be the set $\{1,2\}$, we have that $s_t = \{\langle 1,2\rangle, \langle 1,3\rangle, \langle 2,2\rangle\}$. The element $\langle 3,3\rangle \in s$ is omitted since 3 is not a member of t. Composition of s and r gives $s \circ r = \{\langle 1,1\rangle, \langle 2,2\rangle, \langle 2,3\rangle, \langle 3,2\rangle\}$. The image of t under s is exactly $s[\![t]\!] = \{1,2\}$. These are all the elements that members of t get mapped to under s.

We often write things like $f(x) = x^2$. This notation makes sense for a function f- if x is the input we always get exactly one output $f(x)$. We say f is **well-defined** if $x = y \implies f(x) = f(y)$. That is, there is no ambiguity about the output. If f is a relation but not a function, it may be the case that $f(x)$ is not well-defined. For instance, the previous relation s gives outputs 2 and 3 on input 1. The value $s(1)$ is not well-defined.

Proposition 1. *If f is a (nonempty) function with $\operatorname{dom} f = A$, then $f \circ f^{-1}$ is the identity function on A. That is, $f \circ f^{-1}$ is precisely the set $\{\langle a,a\rangle : a \in A\}$.*

Proof. If $a \in A$ then $\langle a,w\rangle \in f$ for some w in $\operatorname{ran} f$. Therefore $\langle w,a\rangle \in f$. Composing with f^{-1} gives $\langle a,a\rangle \in f \circ f^{-1}$. There are no distinct elements a, b such that $\langle a,b\rangle \in f \circ f^{-1}$, for then this contradicts the definition of f^{-1}. □

Definition 7 (Injective and surjective functions)**.** *Let $f\colon A \to B$ be a function.*

(a) We say that f is injective if $f(x) = f(y)$ implies that $x = y$. Formally, that is to say: $\forall x, y \in A((\langle x,a\rangle = \langle y,a\rangle \wedge \langle x,a\rangle, \langle y,a\rangle \in f) \implies x = y)$. We call f an injection.

(b) We say that f is surjective if for any b in the set B there is a member a in A such that $f(a) = b$. That is, $\forall b \in B(\exists a \in A(\langle a,b\rangle \in f))$. We call f a surjection.

Theorem 2 (Left and right inverses)**.** *Let $f\colon A \to B$ be a function.*

(a) If there exists a function $g\colon B \to A$ such that $g \circ f\colon A \to A$ is the identity function then f is injective.

(b) If there exists a function $h\colon B \to A$ such that $f \circ h\colon B \to B$ is the indentity function then f is surjective.

If they exist, the functions g and h above are called the left and right inverses of f.

Proof. (a) Suppose that g is a map from B to A such that $g \circ f$ is the identity function and suppose there exists $x, y \in A$ such that $f(x) = f(y)$. As $g \circ f$ is the identity function on A then $g(f(x)) = x$ and $g(f(y)) = y$. By assumption g is a function, so since $f(x) = f(y)$ then $g(f(x)) = g(f(y))$. That is, $x = y$, and so f is injective because x, y were arbitrary.

(b) Let h be a function from B to A such that $f \circ h$ is the identity map on B. Let $b \in B$ arbitrary. By assumption, $f(h(b)) = b$. But $\operatorname{dom} h = B$, thus $h(b)$ is a member of A. Hence $h(b)$ lies in the domain of f, which is A, so we find that f maps $h(b)$ to b. As b was arbitrary then $B = \operatorname{ran} f$, and so f is a surjection.

□

In fact, the statements above can be made into if and only if statements under a new assumption. If we have an arbitrary relation R, we've seen that R need not be a function. For instance, $a \in \operatorname{dom} R$ may be related to both b and c in $\operatorname{ran} R$. Can we choose one of these values of b and c to represent the value of a new function that preserves the relation of R?

Axiom 7 (Axiom of Choice, version one)**.** *For any relation R there exists a function $G \subseteq R$ with $\operatorname{dom} G = \operatorname{dom} R$. That is, $\forall A, B(\forall R \subseteq A \times B(\exists G \subseteq R(\operatorname{dom} G = \operatorname{dom} R \wedge \forall x \in \operatorname{dom} G(\exists z \in B(\langle x, z\rangle \in G \wedge (\langle x, z\rangle, \langle x, w\rangle \in G \implies z = w))))))$.*

The axiom of choice gets its name from the nature of what it tells us. Given any relation, the reader has a choice of what function to make. The axiom of choice doesn't not tell us what this function is, but it says there is an existence of such a function. Under the axiom of choice, we can show that every surjection has a right inverse. If $f\colon A \to B$ is a surjection and $b \in B$ is arbitrary, then the preimage $f^{-1}[\![b]\!]$ is nonempty. Consider the relation $r \subseteq B \times A$ where $xry \iff f(y) = x$. For any b as above, there exists a member $a \in A$ where bra. That is, $f(a) = b$. Let $g \subseteq r$ be a function given by the axiom of choice. Then $f \circ g$ is the identity map on B, and g is a right inverse of f.

2.5 Equivalence Relations

We can extrapolate some nice properties out of relations that may make them easier to work with.

Definition 8 (Reflexive, Symmetric, Transitive)**.** *Let R be a relation with a field A (that is, $\operatorname{fld} R = A$.) We say that* *(a) R is reflexive if for all $x \in A$, we have xRx ($\langle x, x\rangle \in R$.)*

(b) R is symmetric if for any $x, y \in A$, we have $xRy \iff yRx$ (if $\langle x, y\rangle \in R$ then $\langle y, x\rangle \in R$ and vice versa.)

(c) R is transitive if for any $x, y, z \in A$ where xRy and yRz, we have xRz ($\langle x, y\rangle \in R, \langle y, z\rangle \in R \implies \langle x, z\rangle \in R$.)

Definition 9 (Equivalence Relation)**.** *We say that a relation R is an equivalence relation if R is reflexive, symmetric and transitive.*

Not all relations are reflexive, symmetric or transitive. The ones that do satisfy these properties are special, and we denote them as **equivalence relations**. Consider the vector space $\mathbb{R}^n$ where n is a natural number. The relation R defined by $uRv \iff u = \lambda v$ for some nonzero scalar λ is an equivalence relation. For any $u \in \mathbb{R}^n$ we have $u = 1 \cdot u$, hence uRu so R is reflexive. If uRv by $u = \lambda v$, then we have vRu via $v = \frac{1}{\lambda}v$, which is valid since λ is nonzero. Hence R is symmetric. If uRv and vRw then $u = \lambda v$ and $v = \nu w$ for some λ, ν. Therefore uRw as $u = \lambda v = \lambda(\nu w) = (\lambda \cdot \nu)w$. Therefore R is also transitive. Indeed, R is an equivalence relation.

In abstract algebra, you may have seen the use of equivalence classes when defining the quotient group G/N where $N \trianglelefteq G$ is a normal subgroup. The **equivalence class** of an element x under an equivalence relation $R \subseteq X \times Y$ is defined as the set $\{y \in A \colon xRy\}$ and is denoted by $[x]_R$, where the subscript $_R$ is often ommitted. Note that the equivalence class of any element x contains x itself, since xRx as R is reflexive. Moreover, if $y \in [x]_R$ then $[x]_R = [y]_R$ by symmetry and transitivity. All elements in the same equivalence class belong to that unique equivalence class. That is to say, the equivalence classes *partition* the set X.

Definition 10 (Partition)**.** *A partition of the set A is a set $\Pi \subseteq \mathcal{P}(A)$ such that*

1. Any two sets in Π are disjoint. That is, $x, y \in \Pi \implies x \cap y = \varnothing$.

2. All members of A are in the union of Π. By that, we mean $\bigcup \Pi = A$.

Definition 11 (Quotient)**.** *Let R be an equivalence relation with $\operatorname{dom} R = X$. The quotient of X modulo R is given by $A/R \coloneqq \{[x]_R \colon x \in X\}$, the set of all equivalence classes under R.*

Theorem 3. *Let R be an equivalence relation on a set X. Then X/R is a partition of X. Moreover, if Π is another partition of X then*

$$S \coloneqq \{\langle x, y\rangle \in X \times X \colon \exists B \in \Pi(x, y \in B)\} \subseteq X \times X$$

is also an equivalence relation on X.

Proof. Let R be an equivalence relation where $\operatorname{fld} R = X$ and consider the quotient X/R. Suppose $[a]_R$ and $[b]_R$ are unique members of X/R. If $[a]_R \cap [b]_R \neq \varnothing$ then there is an element c such that cRa and cRb. Since R is an equivalence relation then aRc by symmetry and hence aRb by transitivity. That is, $[a]_R = [b]_R$, a contradiction. It is clear that $\bigcup X/R \subseteq X$. On the other hand, if $a \in X$ then $[a]_R \in X/R$ and hence $a \in \bigcup X/R$. Thus any sets in X/R are disjoint and the union of X/R is X itself. Indeed, X/R is a partition of X.

Now, if Π is another partition then we'd like to leverage this fact to show that S from above is an equivalence relation on X. For reflexivity, suppose $x \in X$. As Π partitions X, there must be a set $B \in \Pi$ such that $x \in B \in \Pi$. Therefore, $\langle x, x\rangle \in S \implies xSx$. Indeed, if xSy then also ySx by definition. So S is symmetric. Finally, if xSy and ySz, then $x, y \in B$ and $y, z \in C$ for some sets $B, C \in \Pi$. But Π is a partition, so $B = C$ as $B \cap C \neq \varnothing$ (two sets in a partition must be disjoint.) Thus x, z lie in $B \in \Pi$, hence $\langle x, z\rangle \in S \implies xSz$. That is, S is also transitive. □

Consider the relation $R \subseteq \mathbb{N}^2$ where $(x_0, x_1)R(y_0, y_1) \iff x + y_1 = y_0 + x_1$. For example, $(1, 0)R(3, 2)$ since $1 + 2 = 3 = 0 + 3$. Moreover, $(1, 0)$ is not related to $(0, 0)$ since $1 + 0 \neq 0 + 0$. In general, $(m, n)R(m + d, n + d)$ for any naturals m, n and d. As a matter of fact, $\mathbb{R}$ is an equivalence relation (more on this later). The quotient $\mathbb{N}^2/R$ is the set $\{[(0, 0)]_R, [(1, 0)]_R, [(0, 1)]_R, [(2, 0)]_R, \ldots\}$. What is this structure?

Chapter 3

Number Systems

If you've read been reading closely and thinking like a logician, you may have noticed some details that were omitted. For instance, how did we define the Cartesian product $\mathbb{N} \times \{0, 1\}$ in section 2.2 if we did not define what the set $\mathbb{N}$ is? Moreover, $\mathbb{N}$ is an infinite set. With the axioms we have so far, how can we obtain an infinite set? After all, the power set and pairing axioms only generate finite sets (right?) In this section, we dive into the world of infinity (and beyond), as we tackle the question of carefully defining our beloved number systems: $\mathbb{N}, \mathbb{Z}, \mathbb{Q}$ and $\mathbb{R}$.

3.1 The Natural Numbers

The set $\mathbb{N}$ gets its name from the fact that a natural number $n \in \mathbb{N}$ is quite literally, very natural. For instance, if I ask you to buy three apples at the store for me, you understand what to do. I could be greedy and ask for 51 apples, and if I'm a lunatic who cannot get enough I may ask you to handpick 10^{26} beautiful honeycrisps from Zeus's garden. In any case, the number I give you is natural- you can theoretically count it out. On the other hand, the number $\pi = 3.1415926\ldots$ is- to say the least- not natural. How do we list all infinite digits of π? Is it actually *real*? If we interpret numbers as sets, we can be precise.

Definition 12 (Successor, Inductive sets). *For any set a, the successor of a is given by $a \cup \{a\}$, and is denoted as a^+. A set A is inductive if $\varnothing \in A$ and A is closed under successor (that is, for any $a \in A$ we have $a^+ \in A$.)*

A-ha, a plus sign! Surely that seems familiar. Could it be that the successor is just the number plus 1? I mean, $0 + 1 = 1$, but we haven't quite defined addition yet. Moreover, must every inductive set be infinite? If $a \in A$ and A is inductive then $a^+ \in A$. But again, this means $(a^+)^+ \in A$, and so on. Here, we make a leap of faith by placing another (crucial) axiom.

Axiom 8 (Axiom of Infinity)**.** *There exists an inductive set. In first order logic, this axiom says:* $\exists I(\forall x(x \in I \implies (x \cup \{x\} \in I)))$*.*

Again, let's be clear that we are just *assuming* that such a set I exists. To this end, let S be an inductive set. Taking 0 to mean $\varnothing$, we have that $0 \in S$ by definition of being inductive. But indeed, this means $0^+ = \varnothing \cup \{\varnothing\} = \{\varnothing\}$ is a member of S. Denote 1 to be 0^+. Continuing, we have $2 = 1^+ = \{\varnothing\} \cup \{\{\varnothing\}\} = \{\varnothing, \{\varnothing\}\}$ is a member of S, and so on. For any **natural number** $n > 0$, we have $n = \{0, 1, \ldots, n-1\}$. The successor is given by $n^+ = n + 1$ and is the set $\{0, 1, \ldots, n-1, n\}$. The set describing n has n members and contains all sets $m < n$.

Definition 13 (Natural Numbers)**.** *A set is a natural number if it belongs to all other inductive sets. The set of all natural numbers are denoted by* $\mathbb{N}$*, and is the smallest inductive set. That is,* $\mathbb{N}$ *is the intersection of all other inductive sets (we sometimes write* $\mathbb{N}$ *as* ω*, as we will see later on.)*

Proposition 2 (Induction Principle)**.** *If* $A \subseteq \omega$ *is inductive then* $A = \omega$*.*

Proof. If A is inductive then $0 \in A$. Indeed, $0 \in \omega$ as well. Since A is inductive then $0^+ = 1 \in A$. Thus $1^+ = 2 \in A$, and so on. Any natural number $n \in \omega$ must also be a member of A, hence $\omega \subseteq A$. Thus $A = \omega$ as A is already a subset by assumption. □

The proof of the Induction Principle may seem uninteresting, but it is rather useful for everyday mathematics. In mathematical induction, we often show that a statement P holds on some base case (say $P(0)$ is true.) If we let k be arbitrary and assume $P(k)$ holds, then $P(k) \implies P(k+1)$ being true means that $P(n)$ holds for any natural number n. The following Theorem states a trivial fact, but the logic used in its proof is crucial to getting a good grasp at the methods needed for this subject.

Theorem 4. *Every natural number except zero is a successor.*

Proof. Define the set $A \coloneqq \{n \in \omega \colon n = 0 \vee \exists x(x^+ = n)\}$. This set exists by the subset axiom since we have a well-formed formula. We see that A is inductive. For we have $0 \in A$ and if $x \in A$, then x^+ is a successor. If $n = x^+$ then there exists x such that $x^+ = n$. Thus,

$n = x^+ \in A$. That is, A is closed under successor so it is inductive, thus $\omega \subseteq A$. But we defined A as a subset of ω, so by the Induction Principle we have that $A = \omega$. That is, A is the set of natural numbers. In other words, any natural number n is either zero or it is the successor of some set. $\square$