Set Theory

Patrick Oare

1 Set Theoretic Preliminaries

Before diving into any material, we'll begin with preliminary concepts that are found in essentially any discussion of mathematics. There will probably be some concepts that I miss discussing here, in which case I will update this section later. This section is based on:

1. Chapter 0.0 (Basics) of *Abstract Algebra* by Dummit and Foote.
2. Chapter 2.1 (Maps) of *Geometry, Topology, and Physics* by Nakahara.
3. Chapters 1, 2, 3, 6 of *Elements of Set Theory* by Enderton.

The first four subsections (Sets and Quantifiers, Maps, Equivalence Relations, and Countability) are important preliminary material that we will go through before moving onto more interesting topics. I will be adding additional chapters after this which discuss set theory as a whole. For a sneak peak of these additional sections (1.5 onward), check out Fig. 1.

1.1 Sets and Quantifiers

The most fundamental object in mathematics is the **set**. A set is essentially a container that contains objects. To denote that an element x is in a set A, we use the notation

$$x \in A. \tag{1}$$

We denote sets with curly braces, and instantiate them in two ways. The first is to simply write out the elements that a set contains, i.e. $A_0 = \{1, 3, 5, 7, 9\}$. This is practical to do when the set is small, but may be difficult when the set is large. In this case, we use *set builder notation*, in which we specify another set to sample from, given some conditions. For example, the previous set can be constructed as,

$$A_0 = \{x \in \mathbb{Z} : 1 \leq x \leq 10 \text{ and } x \text{ is odd}\}. \tag{2}$$

Here $\mathbb{Z}$ denotes the integers. This reads "the set of x in the integers such that $1 \leq x \leq 10$ and x is odd", which is of course the set $\{1, 3, 5, 7, 9\}$. The symbol :, and also |, means **such that** (this is overloaded; one will often see either a colon or a vertical bar used in set builder notation). The **empty set** (the set containing no elements) is denoted by

$$\emptyset := \{\}. \tag{3}$$

TWO

What is a two? What are numbers? These are awkward questions; yet when we discuss numbers one might naively expect us to know what it is we are talking about.

In the Real World, we do not encounter (directly) abstract objects such as numbers. Instead we find physical objects: a number of similar apples, a ruler, partially filled containers (Fig. 26).

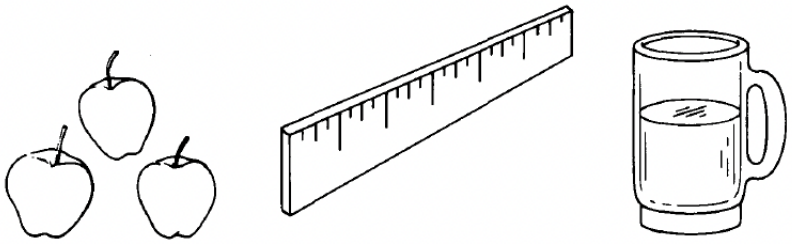

Fig. 26. A picture of the Real World.

Somehow we manage to abstract from this physical environment the concept of numbers. Not in any precise sense, of course, but we feel inwardly that we know what numbers are, or at least some numbers like 2 and 3. And we have various mental images that we use when thinking about numbers (Fig. 27).

Figure 1: Chapter 6 of Enderton's *Elements of Set Theory*, my favorite chapter from any math book. Most of the math we will do in this course is practical, and we will be learning it with an eye towards actual applications. However, set theory by itself is a much more philosophical branch of math. It asks how we construct the objects we work with. We usually just assume that numbers exist and go from there, but in order to be rigorous, we must carefully define anything we want to work with. This area of math is filled with lots of pedantics, but also offers deep insight into things like the nature of infinity.

The := symbol is **definition**. A set B is a **subset** of A if each element in B is contained in A, and denote this as $B \subseteq A$ (i.e., $\{1,5\} \subseteq A_0$ in the previous example). Every non-empty set A has two distinct subsets, $\emptyset \subseteq A$ and $A \subseteq A$. We use the symbol $\subset$ to denote a proper subset, which means that the sets are not equal, i.e. $B \subset A$ if $B \subseteq A$ and $B \neq A$. Note that two sets A and B are equal iff they are subsets of one other:

$$A = B \iff A \subseteq B \text{ and } B \subseteq A. \tag{4}$$

Often in proofs, the easiest way to show sets are equal is to prove both of these containments.

One can form a set from all the subsets of another set A: this set is called the **power set** of A, denoted 2^A. The reason it is called the power set is because if A is finite, then 2^A is also finite, with cardinality $|2^A| = 2^{|A|}$. For example, if $A = \{3, 9\}$, then

$$2^A = \{\emptyset, \{3\}, \{9\}, \{3, 9\}\}. \tag{5}$$

As in the example above, the power set will always contain the empty set $\emptyset$ and the set A itself.

We will often use **quantifiers** to concisely express mathematical ideas. The quantifier $\forall$ means "for all", and the quantifier $\exists$ means "there exists". These will be our default way of making mathematical statements. For example, the statement,

$$\forall n \in N, n < n + 1, \tag{6}$$

is a **tautology** (a statement that is always true), and means "each natural number n is less than itself plus 1". We will see how these are used in some simple examples below.

There are a number of important operations that we can do on sets. The size of a set is called its **cardinality**, denoted card(A), or simply $|A|$. The cardinal of the set $A_0 = \{1, 3, 5, 7, 9\}$ in the previous example is 5, because it has five elements. The cardinal $\text{card}(\emptyset) = 0$, which is its defining characteristic. The empty set is the basic building block of mathematics; if you're interested in how the natural numbers are properly constructed starting from $\emptyset$, check Interlude 1. The **union** of two sets A and B, $A \cup B$, is the set of elements contained in either A and B,

$$A \cup B = \{x : x \in A \text{ or } x \in B\}. \tag{7}$$

The **intersection** of A and B is the set of elements contained in both A and B,

$$A \cap B = \{x \in A : x \in B\} = \{x \in B : x \in A\}. \tag{8}$$

Given a collection of sets $\{A_i\}_{i=1}^K$ for some $K \in \mathbb{N} \cup \{\infty\}$, we can form a union or intersection of the entire collection of sets, denoted with a large union or intersection symbol,

$$\bigcup_{i=1}^{K} A_i = \{x : \exists i \text{ s.t. } x \in A_i\} \qquad \bigcap_{i=1}^{K} A_i = \{x \in A_1 : \forall i,\, x \in A_i\}. \tag{9}$$

We will also consider the **set difference** of two sets. The set difference of two sets A and B is the collection of elements of A which are not in B,

$$A \setminus B = \{x \in A : x \notin B\}. \tag{10}$$

If $B \subseteq A$, the **complement** of B in A, denoted B^c, is

$$B^c = A \setminus B. \tag{11}$$

The complement of B is everything that is in its external universe (A) that is not in B.

How large are these sets? The union of two sets adds together the elements of each set, minus the number of elements in the intersection,

$$\left|A \cup B\right| = \left|A\right| + \left|B\right| - \left|A \cap B\right|. \tag{12}$$

This idea generalizes for arbitrary unions to the inclusion-exclusion principle.

Proposition 1: Inclusion-Exclusion Principle

For sets $\{A_i\}_{i=1}^N$, we have:

$$\left|\bigcup_{i=1}^N A_i\right| = \sum_{k=1}^N (-1)^{k+1} \sum_{1 \le i_1 \le \ldots \le i_k \le N} \left|A_{i_1} \cap \ldots \cap A_{i_k}\right|. \tag{13}$$

I don't particularly care about this formula, but for some reason lots of math professors spend way too much time on it, so you do you if you want to study the proof or not. For $A \setminus B$, the suggestive name of "set difference" indeed applies to the size of the set difference,

$$|A \setminus B| = |A| - |B|. \tag{14}$$

Two sets are **disjoint** if they share no elements: $A \cap B = \emptyset$. This definition extends to collections of sets. Suppose we have a collection of sets $\{A_i\}_{i \in I}$, where I is an index set[1]. We say that $\{A_i\}_{i \in I}$ is disjoint if this collection of sets is **pairwise disjoint**; that is, if

$$A_i \cap A_j = \emptyset \qquad \forall i \neq q \in I. \tag{15}$$

When two sets are disjoint, we often denote their union as $A \sqcup B$, and call it the **disjoint union**. It is exactly the same operation as the regular union, just the $\sqcup$ specifies that the two sets are disjoint. We likewise can take the disjoint union of a collection of sets,

$$\bigsqcup_{i \in I} A_i. \tag{16}$$

Note that the cardinality of disjoint unions adds,

$$\left|\bigsqcup_{i \in I} A_i\right| = \sum_{i \in I} |A_i|, \tag{17}$$

[1]Collections of sets are often written like this. I is called the **index set**, and often taken to be $\mathbb{N}$ or some finite subset of $\mathbb{N}$. For example, if $I = \{0, 1, 2, 3, 4\}$, then we are looking at the collection $\{A_i\}_{i \in I} = \{A_0, A_1, A_2, A_3, A_4\}$. If $I = \mathbb{N}$, then we are looking at the collection $\{A_i\}_{i \in I} = \{A_i\}_{i=0}^{\infty}$. We'll see more complicated index sets as we go along in this course.

because we can ignore all the intersection terms in the inclusion-exclusion principle (Prop 1).

Finally, we consider products of sets. First, the **Cartesian product**. This is defined as the set of **ordered pairs** (a, b), where $a \in A$ and $b \in B$,

$$A \times B := \{(a, b) : a \in A, b \in B\}. \tag{18}$$

The reason for the suggestive notation is because for finite sets, the cardinality multiplies:

$$\big|A \times B\big| = \big|A\big| \times \big|B\big| \tag{19}$$

This is the second way we have found to join A together with B; the first way was the union. Note that for disjoint sets $A \cap B = \emptyset$, the union adds together their cardinals, while the Cartesian product multiplies the cardinals. The Cartesian product should be thought of as "multiplying" sets together, while the union should be thought of as "adding" sets together. We can likewise generalize the product of sets to arbitrary collections with $\prod_{i \in I} A_i$.

Interlude 1: Ordered pairs

Sets do not care about the ordering of their elements: $\{x, y\} = \{y, x\}$, and so on for larger sets. In contrast, ordered pairs are objects that can be constructed via set theory, with the fundamental definition that $(x, y) = (a, b)$ if and only if $x = a$ and $y = b$. Note that this is distinct from the set $\{x, y\}$, as, for example, $(1, 2) \neq (2, 1)$ but $\{1, 2\} = \{2, 1\}$, as sets are unordered. One can construct the ordered pair from set theory: the definition used frequently today is the **Kuratowski definition** which defines,

$$(x, y) := \{\{x\}, \{x, y\}\}. \tag{20}$$

One can prove that this definition works: that $(x, y) = (u, v)$ if and only if $x = u$ and $y = v$. Clearly $x = u, y = v$ implies that $(x, y) = (u, v)$, so we only need to prove the converse. Suppose $(x, y) = (u, v)$ with this definition, so $\{\{x\}, \{x, y\}\} = \{\{u\}, \{u, v\}\}$. This gives two possibilities:

1. $\{x\} = \{u, v\}$ and $\{x, y\} = \{u\}$. In this case, since we have a set of cardinal 1 equalling a set of cardinal 2, this implies that $u = v$ and $x = y$. But then, $\{x\} = \{u, v\} = \{u\}$ (since repeated elements are thrown out of a set) implies $x = u$, and likewise $y = x = u = v$.

2. $\{x\} = \{u\}$ and $\{x, y\} = \{u, v\}$. In this case, we see $x = u$ and $y = v$ as desired.

We see that in either case, we have the definition we wanted!

To extend this definition to **ordered n-tuples**, we want the Cartesian product definition of sets to be associative, i.e., $A^3 = (A \times A) \times A$. This means the ordered triple should be $(x, y, z) := ((x, y), z)$, and in general extend this definition inductively,

$$(x_1, ..., x_n) := ((x_1, ..., x_{n-1}), x_n). \tag{21}$$

Note that this is **not** the naïve Kuratowski definition we would obtain from $(x, y, z) = \{\{x\}, \{x, y\}, \{x, y, z\}\}$. Why does this not work? (Hint: consider using this definition on the triples $(1, 2, 1)$ and $(1, 1, 2)$)

Exercise 1

Are any of the sets $\emptyset$, $\{\emptyset\}$, or $2^{\emptyset}$ equal? What are the cardinalities of each set? What about $2^{2^{\emptyset}}$?

A number of sets will frequently come up in our discussions. We will assume each of these sets exists; in formal set theory, one must axiomatically state the properties that the sets have, then *construct* an explicit system that satisfies these axioms[2]. The **natural numbers** are denoted by $\mathbb{N}$,

$$\mathbb{N}_0 = \{0, 1, 2, 3, ...\}. \tag{22}$$

See Interlude 1 for a further discussion of the natural numbers. We will often use $\mathbb{N}$ to denote counting numbers, $\mathbb{N} = \{1, 2, 3, ...\}$, and include the 0 subscript to denote we are also including 0. The **integers** are denoted $\mathbb{Z}$,

$$\mathbb{Z} = \{..., -2, -1, 0, 1, 2, ...\}. \tag{23}$$

The **rational numbers** $\mathbb{Q}$ are the set of fractions a/b, where a and b are integers,

$$\mathbb{Q} = \left\{ \frac{a}{b} : a \in \mathbb{Z}, b \in \mathbb{Z} \setminus \{0\} \right\}. \tag{24}$$

The **real numbers** are denoted by $\mathbb{R}$. I can't give an explicit set construction of them without going into Dedekind cuts or convergent sequences, but just know that $\mathbb{R}$ "fills in the gaps" of the rational numbers. We'll discuss this further in Chapter **??**. Finally, the **complex numbers** $\mathbb{C}$ are defined as formal pairs of real numbers,

$$\mathbb{C} = \{a + ib : a, b \in \mathbb{R}\}. \tag{25}$$

One has the sequence

$$\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C}. \tag{26}$$

As we go up in grade and consider larger sets, we inherit more structure. For example, only addition is defined on $\mathbb{N}$, but addition, subtraction, and multiplication are defined on $\mathbb{Z}$. Division is defined in $\mathbb{Q}$, and $\mathbb{R}$ and $\mathbb{C}$ have even more specialized structures that we'll discuss later. The interplay of these operations $(+, -, \times, \div)$ with different structures will be considered when we discuss algebra. Finally, given $n \in \mathbb{N}_{>0}$, we'll use the Cartesian product to represent the product of n copies of the same set. For example,

$$\mathbb{Z}^3 = \mathbb{Z} \times \mathbb{Z} \times \mathbb{Z} = \{(n, m, p) : n, m, p \in \mathbb{Z}\}. \tag{27}$$

The most important application of this is **Euclidean space**, which is the set $\mathbb{R}^n = \mathbb{R} \times ... \times \mathbb{R}$.

[2]The constructions of these sets are often quite interesting, and we may explicitly construct $\mathbb{R}$ and $\mathbb{C}$ later in these notes, because those in particular are quite illuminating as to the structure of the underlying sets. An example of a set construction is the construction of the natural numbers $\mathbb{N}$, shown in Interlude 1.

1.2 Maps (aka functions)

A **map** between sets is simply a function $f : A \to B$ which assigns an element of B to each element of a. We denote this by $f(a) \in B$, for each $a \in A$. We often will use shorthand notation and denote the map f as follows when the context is clear,

$$\begin{aligned} A &\to B \\ a &\mapsto f(a). \end{aligned} \tag{28}$$

The set A is called the **domain** of f, and the set B is the **codomain** of f. The **image** of f is the set of values in the codomain that it takes on,

$$\mathrm{im}(f) := \{b \in B : \exists a \in A \text{ s.t. } b = f(a)\} \subseteq B. \tag{29}$$

The image is a subset of the codomain, but does not necessarily have to be the entire codomain. The codomain is the set of values that the map is *allowed* to take on, but the image is the set of values that the map *does* take on. Consider the map $f_3 : \mathbb{N} \to \mathbb{N}$, $n \mapsto n + 3$. $\mathbb{N}$ is both the domain and the codomain of the map, but the image of f_3 is $\{3, 4, 5, ...\}$, because there is no natural number such that $0 = f_3(n)$, or $1 = f_3(n)$, or $2 = f_3(n)$.

Given two functions $f : A \to B$ and $g : B \to C$, we can **compose** the functions to create a map $g \circ f : A \to C$, defined by,

$$a \mapsto g(f(a)) \in C. \tag{30}$$

Note that we read function composition from right to left. This can often get confusing when things get complicated, so we will often use **commutative diagrams** to show function composition. In my opinion, commutative diagrams make things so much clearer. For example, the previous composition $g \circ f : A \to C$ is denoted by the diagram,

$$\begin{array}{ccccc} A & \xrightarrow{f} & B & \xrightarrow{g} & C \\ & \searrow & & \nearrow & \\ & & g\circ f & & \end{array} \tag{31}$$

Another operation on functions that will prove important is restriction to a subset of the domain. Suppose $f : A \to B$, and $Z \subseteq A$ is a subset. The **restriction** of f to Z, denoted $f|_Z$ or $f|Z$, is the map f with the domain Z,

$$f|_Z : Z \to B \qquad\qquad f|_Z(x) = f(x) \tag{32}$$

For example, consider the map $g_3 : \mathbb{Z} \to \mathbb{Z}$, $x \mapsto x + 3$. This has the same formula as f_3, but the domain is different. The restriction $g_3|\mathbb{N}$ is *almost* the same as f_3, but the codomain is different. However, note that $\mathrm{im}(g_3|\mathbb{N}) = \mathrm{im}(f_3) = \{3, 4, 5, ...\}$. Because $\mathrm{im}(g_3|\mathbb{N}) \subseteq \mathbb{N}$, we can equally well restrict the codomain of $g_3|\mathbb{N}$ to be $\mathbb{N}$, in which case this restriction map equals f_3. A lot of this probably feels like semantics, but it can be quite important in specific contexts.

A map is called **surjective**, or **onto**, if its image equals its codomain. If a map f is surjective, we will often denote it by

$$x \twoheadrightarrow f(x). \tag{33}$$

A map is **injective**, or **one-to-one**, if each element in the domain maps to a unique element in the codomain. That is, f is injective is

$$\forall x, y \in A, f(x) = f(y) \implies x = y. \tag{34}$$

Here the $\implies$ arrow means **implies**. If f is injective, we denote it as

$$x \hookrightarrow f(x). \tag{35}$$

We will later call specific types of injective maps **embeddings**, for reasons that we will soon see. A map which is both injective *and* surjective is called a **bijection**, denoted

$$x \xrightarrow{\sim} f(x). \tag{36}$$

A bijective map $f : A \to B$ has an **inverse** $f^{-1} : B \to A$ that satisfies $f \circ f^{-1} = \mathrm{id}_B$ and $f^{-1} \circ f = \mathrm{id}_A$, where id_S denotes the identity map on the set S. This is represented by the **commutative diagrams**,

$$\begin{array}{ccc} A & \xrightarrow{f} & B \\ & \searrow^{id_A} & \downarrow{\scriptstyle f^{-1}} \\ & & A \end{array} \qquad \begin{array}{ccc} B & \xrightarrow{f^{-1}} & A \\ & \searrow^{id_B} & \downarrow{\scriptstyle f} \\ & & B \end{array} \tag{37}$$

If a bijection $f : A \xrightarrow{\sim} B$ exists between A and B, it means that they have the same number of elements, $\mathrm{card}(A) = \mathrm{card}(B)$. Strictly as sets, we can consider A and B to have the same structure, just as relabelings of one another. For example, the sets $A = \{a, b, c\}$ and $B = \{1, 2, 3\}$ are in bijection with one another. They are each 3 element sets, and as sets function in the same way[3].

One reason that injective maps are particularly important is because they **induce** a bijection onto their image. Suppose that $f : A \hookrightarrow B$ is an injection. We can consider the map f to have codomain $\mathrm{im}(f)$ without loss of generality (WLOG). With this smaller codomain, the map becomes a bijection,

$$f : A \xrightarrow{\sim} \mathrm{im}(f) \subseteq B, \tag{38}$$

and we can construct an inverse map $f^{-1} : \mathrm{im}(f) \to A$. We will often call this type of map an embedding, since we can view f as "embedding" A into B, as A is equivalent to $\mathrm{im}(A)$, which is a subset of B[4].

[3]You can argue that "I can add together elements of B, but not elements of A". But, the addition operation is additional structure that we add onto B once we semantically know that B contains numbers, not arbitrary symbols; it's not a native part of the set B constructed as $\{1, 2, 3\}$, so as sets, A and B are equivalent. We'll see later that B inherits an addition operation from A because of the map f; this is called a pull-back, and we'll see this basic idea in lots of different contexts to be discussed later

[4]These definitions will slightly change when A and B have additional structure, which we will see in the next chapter.

1.3 Equivalence Relations

Equivalence relations allow us to create new sets from old ones by "identifying" different points in the old set as equal. It imposes a new structure that allows us to generate smaller objects from bigger objects. As we'll see in the section on algebra, this is the formalism in which quotient objects will be defined, which are hugely important in all areas of math.

A **relation** between sets is a pairing between the sets that is either true or false. Examples of relations on the natural numbers are $\leq$, $=$, and $\geq$. The formal definition of a relation is below, which we will explore next.

Definition 1: Relation, Equivalence Relation

A binary relation R on a set A is a subset of $A \times A$,

$$R \subseteq A \times A. \tag{39}$$

We write $a \sim b$ if $(a, b) \in R$. We say R is an **equivalence relation** if:

1. ($\sim$ is **reflexive**) For each $a \in A$, $a \sim a$.
2. ($\sim$ is **symmetric**) If $a \sim b$, then $b \sim a$.
3. ($\sim$ is **transitive**) If $a \sim b$ and $b \sim c$, then $a \sim c$.

Relations allow one to generalize symbols like $\leq$, $=$, and $\geq$ to other settings; equivalence relations are even more stringent, and can take the form of $=$ on other sets. Consider these relations on the set $A = \{1, 2, 3\}$. These relations can be represented as a subset of $A \times A$ (as pairs of elements of A) as:

$$\begin{aligned} R_{\leq} &:= \{(1,1), (1,2), (1,3), (2,2), (2,3), (3,3)\} \\ R_{=} &:= \{(1,1), (2,2), (3,3)\} \\ R_{\geq} &:= \{(1,1), (2,1), (3,1), (2,2), (3,2), (3,3)\}. \end{aligned} \tag{40}$$

Play with these examples and make sure they make sense. For example, saying $1 \leq 2$ on A is just saying that the pair $(1, 2)$ is an element of $R_{\leq}$. We see that $\leq$, $=$, and $\geq$ are all reflexive and transitive, but *only* $=$ *is symmetric*, meaning that out of these three, $=$ is the only equivalence relation.

The importance of equivalence relations stems from the desire to identify different objects as the same. It is often useful to be able to study objects with a different level of granularity than originally presented; we will see this more prevalently when we look at quotient groups and spaces, and quotient topologies. Given an equivalence relation $\sim$ and an element $a \in A$, one can form the equivalence class of a, which is the set of elements that are equal to a under $\sim$.

Definition 2: Equivalence Classes

Let $\sim$ be an equivalence relation. The **equivalence class** of $a \in A$ is the set of elements

which are related to a under $\sim$, denoted $[a]$,

$$[a] := \{b \in A : b \sim a\}. \tag{41}$$

We also write $[a]$ as $a \mod \sim$. The set of equivalence classes of A is denoted by $A/\sim$ ("A quotient $\sim$").

Note that we can equivalently write $a \sim b$, $a \in [b]$, or $b \in [a]$. We care about equivalence classes because they allow one to partition A into subsets. Essentially, each element of A belongs to an equivalence class (A is the union of all its equivalence classes) each equivalence class is distinct and does not overlap with any other equivalence class. Intuitively, A is "carved up" into equivalence classes. A and $A/\sim$ have the same structure, but $A/\sim$ has a smaller number of elements, as the elements of $A/\sim$ are groups of elements of A.

Definition 3: Partition of a set

Let A be a non-empty set. A partition $\mathcal{P}$ of A is a collection of non-empty subsets of A such that

1. For non-equal $P, Q \in \mathcal{P}$, $P \cap Q = \emptyset$.
2. A is the union of the subsets in $\mathcal{P}$,

$$A = \bigcup_{P \in \mathcal{P}} P. \tag{42}$$

Note that in the notation we've developed thus far, $\mathcal{P} \subseteq 2^A$, i.e. $\mathcal{P}$ is a subset of the power set of A.

We will often write $A = \bigsqcup_{P \in \mathcal{P}} P$, which means that A is the **disjoint union** of the sets in $\mathcal{P}$. Disjoint union, $\sqcup$, is essentially the same thing as a regular union when the individual sets are disjoint (they do not share any elements). Now we prove that the classes of an equivalence relation partition A.

Proposition 2

Let $\sim$ be an equivalence relation on non-empty A. Then the equivalence classes of A, $A/\sim$, are a partition of A.

Proof. Note that each element of $A/\sim$ is non-empty, as for $[a] \in A/\sim$, $a \in [a]$. Suppose that $[a] \neq [b]$ are two distinct equivalence classes. Clearly, $a \not\sim b$, else we would have $[a] = [b]$ by definition. But, we must go one step further; we must show that $[a] \cap [b] = \emptyset$. We argue by contradiction; suppose that $[a] \cap [b] \neq 0$, and let $c \in [a] \cap [b]$. Then $c \sim a$ and $c \sim b$, which by transitivity implies $a \sim b$. But, this is a contradiction to our initial assumption that $a \not\sim b$, hence $[a] \cap [b]$ must be empty. To finish the proof, note that for any $a \in A$, $a \in [a] \in \mathcal{P}$, hence $A \subseteq \mathcal{P}$. The reverse containment is obviously true since $\mathcal{P} \subseteq 2^A$ (each equivalence class is made of elements of A. This establishes $A = \bigsqcup_{P \in \mathcal{P}} P$, i.e. that $\mathcal{P}$ partitions A. $\square$

Let's do a quick example on the set $A = \mathbb{Z}$, where $\sim$ is defined as:

$$n \sim m \text{ iff } n - m = 3k \text{ for some } k \in \mathbb{Z}. \tag{43}$$

This equivalence relation makes two integers equal if they are related by a multiple of 3. For example, $2 \sim 5 \sim 8$ and $1 \sim 31$, but $3 \not\sim 8$. You can verify that this is an equivalence relation. This relation has three equivalence classes; the set $[0] = \{..., -3, 0, 3, 6, ...\}$, the set $[1] = \{..., -2, 1, 4, 7, ...\}$, and the set $[2] = \{..., -1, 2, 5, 8, ..., \}$. Note that we simply picked two *representatives* of these classes; we could equally well call the class [1] by [4] or [121], it doesn't matter as long as it's an integer of the form $1 + 3k$ for some k. This is depicted pictorially in Figure 2.

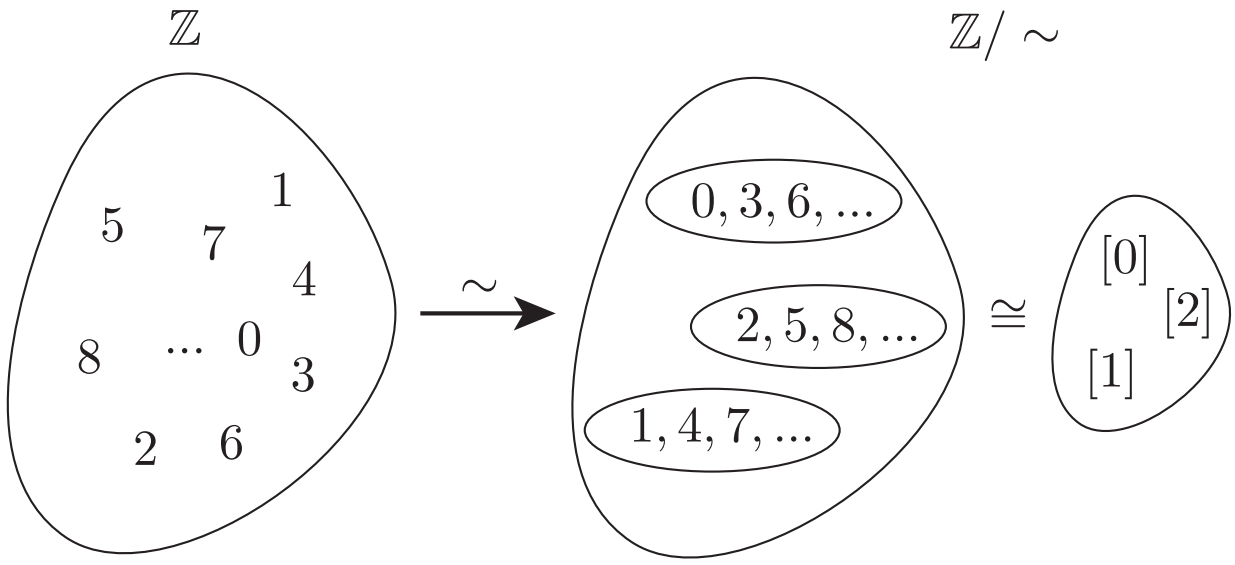

Figure 2: The equivalence relation which relates even and odd numbers to one another, defined in Eq. (43). The set $\mathbb{N}$ is shown on the left as a cartoon, and its equivalence classes are shown on the right. Note that the equivalence classes are simply distinct groupings of elements of $\mathbb{N}$!

The main reason that equivalence classes are useful are that they allow us to define maps and other objects using only the qualities we wish. In this case, we might use the set $\mathbb{Z}/\sim$ if we only care about the value of an integer mod 3. An important idea conveyed in Figure 2 is that although we constructed $\mathbb{Z}/\sim$ as made of equivalence classes (a set of sets), it often functions the same way as a simple set of 3 elements, $\mathbb{Z}/\sim = \{[0], [1], [2]\}$, where now we consider the labels [0], [1], and [2] to be abstract symbols that represent the equivalence classes. Although $\mathbb{Z}/\sim$ has a large amount of structure by construction, it will often function the same way as a set constructed from three elements. We will see this further later when we study quotient groups and have operations that these equivalence classes inherit. The set $\mathbb{Z}/\sim$ will also come back, and we will see it later as the *cyclic group of order 3*, denoted $\mathbb{Z}/3\mathbb{Z}$.

We will often call equivalence classes **cosets**; this will have a more formal definition when we look towards quotient groups. Defining functions on cosets must be done carefully. We've seen that we can represent a coset by one of its elements, which we call a **representative** of the coset. For example, 0 is a representative of [0], but any element in [0] is equally good; the 9 could also have been picked as a representative of the coset, and we could denote it by [9]. This is simply due to the fact that if $n \sim m$, then $[n] = [m]$, so we can use either representative WLOG.

The dangerous thing about representatives is when we define maps on these spaces. We must be sure that all the representatives of a set are sent to the same place. For example, consider the map:

$$\mathbb{Z}/\sim \to \mathbb{Z}/\sim \qquad\qquad [n] \mapsto [n+2] \tag{44}$$

This is defined via a representative of each coset, and so it **must send each possible representative of the coset to the same place**. This map is well-defined. For example, consider where it sends the coset [1] to. It sends $[1] \mapsto [1+2] = [3] = [0]$, and so if we want it to be a well-defined map we need it to send everything in the equivalence class of 1 to [1]. Indeed this does, because any element $n \in [1]$ can be written $n = 1 + 3k$, so $n + 2 = 3 + 3k = 3(k+1) \sim 0$, i.e. $[n+2] = [0]$. This same argument extends to the other equivalence classes to show that addition by a constant is well-defined on each coset.

Interlude 2: Factoring a map through the quotient

Let $f : A \to B$ be a set. We will use equivalence relations to define an injective map $\tilde{f}$ on a quotient of A. The idea here is that for an arbitrary map f, the "non-injectiveness" is a result of different elements in the domain mapping to the same element in the codomain: it comes from $f(x) = f(y)$ for distinct $x \neq y$ values. So, *what if we make all the values in the domain which map to the same point equivalent?* Define a relation $\sim$ on A by,

$$x \sim y \text{ iff } f(x) = f(y). \tag{45}$$

This is an equivalence relation, and you should prove it if you're interested. Now, the quotient of A under $\sim$ is the set of cosets of A under $\sim$,

$$A/\sim = \{[x] : x \in A\}. \tag{46}$$

Each coset $[x]$ contains all the elements $y \in A$ which have the same function value as x, $f(y) = f(x)$. The natural thing to do now is to consider defining a map $\tilde{f}$ **on the quotient** given by

$$\tilde{f} : A/\sim \longrightarrow B \qquad\qquad [x] \mapsto f(x). \tag{47}$$

This map $\tilde{f}$ looks like f: it still has the same image as f, only now it acts on *cosets of* A, not on elements of A. It has grouped up the elements of A into cosets, in which each element in a given coset has the same value of f. Because of that, it is well-defined: if $[x] = [y]$ are represented by the same coset, then $f([x]) = f(x) = f(y) = f([y])$ because $x \sim y$ by definition. This map $\tilde{f}$ is clearly injective, since we've grouped up all the

values in A which equal each other under f into the same coset! So, from f we have constructed an injective map $\tilde{f}$ which acts on the cosets of A under $\sim$: we have used equivalence relations to quotient out all the "non-injectiveness" of the map f.

It is customary to represent this by a commutative diagram. For any quotient of A, there is a canonical map $A \to A/\sim$ that passes an element of A to its corresponding coset,

$$\pi : A \longrightarrow A/\sim \qquad\qquad x \mapsto [x]. \tag{48}$$

Note this map is surjective (its image is every coset of $A/\sim$), and only injective if f itself is. This map is called the **projection to the quotient**: we will see plenty more maps like this later. The map f **factors through the quotient**, in the sense that the following diagram commutes.

$$\begin{array}{ccc} A & \xrightarrow{\pi} & A/\sim \\ & \searrow^{f} & \downarrow{\scriptstyle \tilde{f}} \\ & & B \end{array} \tag{49}$$

This diagram is just a fancy way of saying that

$$f = \tilde{f} \circ \pi. \tag{50}$$

However, writing it out as a commutative diagram provides more information about each individual map, as you can read off the domain and codomain of each map explicitly.

1.4 Countability

The last preliminary concept that we will discuss is countability. If I have time, we'll do a full section later on ordinal numbers, which allow one to rigorously and quantitatively define the concept of infinity. This allows you to also do fun things like add and multiply different types of infinity. For now, we'll just stick to the basics. We begin with the concept of countability.

Definition 4: Countable set

A set is **countable** if it is either finite or in bijection with the natural numbers $\mathbb{N}$. If a set is not countable, we say it is **uncountable**.

In terms of our common sets, $\mathbb{N}$, $\mathbb{Z}$, and $\mathbb{Q}$ are countable, but $\mathbb{R}$ and $\mathbb{C}$ are uncountable. Countability lets us quantify the sizes of infinities that we'll encounter. Even though the relation $\mathbb{N} \subseteq \mathbb{Z}$ makes it seem that $\mathbb{Z}$ is larger than $\mathbb{N}$, they're actually "the same size". The explicit bijection between them is,

$$f : \mathbb{N} \to \mathbb{Z} \qquad\qquad n \mapsto \begin{cases} n//2 & n \text{ is even} \\ -(n+1)//2 & n \text{ is odd} \end{cases}, \tag{51}$$

where // denotes integer division. Note that showing a set is countable is equivalent to **enumerating** the set: that is, showing that you can traverse the set one element at a time and hit every element in the set. The bijection of Eq. (51) corresponds to the enumeration,

$$f \approx 0, -1, 1, -2, 2, -3, 3, ... \tag{52}$$

(the RHS is just $f(0), f(1), f(2), f(3), f(4), f(5), f(6), ...$), which we see surjects $\mathbb{N}$ onto $\mathbb{Z}$. So, you should think of a countable set as *enumerable*: any countable set A can be expressed as $A = (a_i)_{i=1}^{N}$, where $N \in \mathbb{N} \cup \{\infty\}$. In contrast, a set like $\mathbb{R}$ is uncountable and **cannot** be enumerated.

Exercise 2

Show that $\mathbb{Q}$, the set of rational numbers, is countable. *Hint*: Lay the elements of $\mathbb{Q}$ out on a 2-dimensional grid, with 0 at the center (note that each element of the grid does not have to be a unique value of $\mathbb{Q}$, for example we could have the elements $1/2$ and $2/4$, which are the same, on different parts of the grid). Show that you can enumerate the grid.

When proving a set is countable, the order of enumeration is extremely important to specify, and it is often the most important thing to get right in the entire proof. For example, let's take a look at the following theorem.

Proposition 3

The countable union of countable sets is countable.

Proof. Let $\{A_I\}_{i \in I}$ be countable sets A_i indexed by some countable set I (for example, we can think of this like trying to take the union $\bigcup_{i=1}^{\infty} A_i$, where $A_i = \mathbb{N}$). WLOG let $I \subseteq \mathbb{N}$; clearly it suffices to prove the case when $I = \mathbb{N}$, and we are looking at $\{A_i\}_{i=1}^{\infty}$. The point here is to construct an enumeration of $\bigcup_{i=1}^{\infty} A_i$: we have to show that this enumeration reaches every element in the union. We denote by J_i the i^{th} indexing set, so $A_i = (a_j^{(i)})_{j \in J_i}$ with each J_i countable.

The important part of this proof is **how** we enumerate the elements of $\bigcup_{i=1}^{\infty} A_i$. Let $f : \mathbb{N} \to \bigcup_{i=1}^{\infty} A_i$ be the bijection. If we try to enumerate A_1 first before any other A_i, we will not be able to properly define the map; taking $f(i) = a_i^{(1)}$ ends with $\text{im}(f) = A_1$, **not** $\text{im}(f) = \bigcup_{i=1}^{\infty} A_i$. Instead, we must enumerate all these sets jointly; we take $f(1) = a_1^{(1)}$ and $f(2) = a_1^{(2)}$. Crucially, we begin to define the map *diagonally*. We cannot simply traverse the sets with the next elements being $a_1^{(3)}$, $a_1^{(4)}$, $a_1^{(5)}$, and so on: because there are infinite sets, defining the map in this way would never allow us to reach the second element in any of the sets, for example $a_2^{(1)}$. So, we define $f(3) = a_2^{(1)}$. For the third row, we define $f(4) = a_3^{(1)}$, $f(5) = a_2^{(2)}$, and $f(6) = a_1^{(3)}$. This is shown in Fig. 3, and f provides the desired bijection[5]. □

[5]Note that each "row" of the map is defined by $a_i^{(j)}$ with $i + j$ being a constant. Behaviors like this are the hallmarks of diagonal arguments.

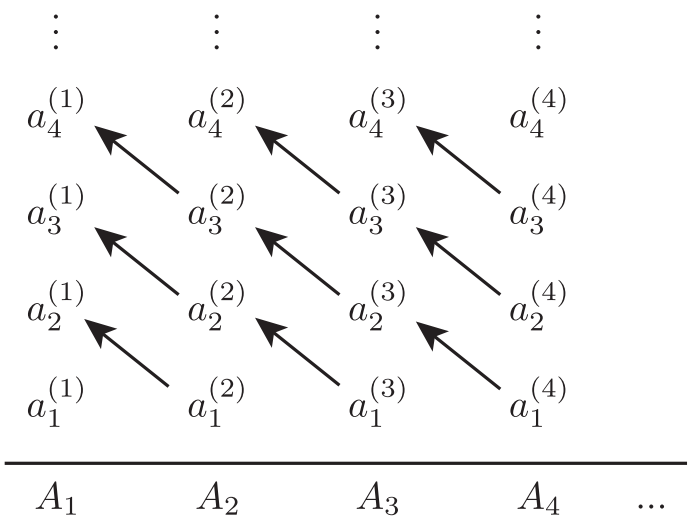

Figure 3: Enumeration of a countable union of a countable set. Note the diagonal argument manifests as diagonal arrows in the grid. The enumeration of $\bigcup_{i=1}^{\infty} A_i$ is defined as $a_1^{(1)}, a_1^{(2)}, a_2^{(1)}, a_1^{(3)}, a_2^{(2)}, a_3^{(1)}, ...$

1.5 What is a number? $\mathbb{N}$, $\mathbb{Z}$, and $\mathbb{Q}$

What the hell is a number? Math is first encountered as the study of numbers and the operations defined on numbers. How do you add 5 and 6? What's 3 times 4? How do you subtract two fractions? We learn about all this in school, and eventually you get really good at manipulating numbers with the basic arithmetic operations: addition, subtraction, multiplication, and division. In my opinion, more than anything else, math is the study of logic, not numbers. Numbers merely provide a stepping stone into understanding the different areas of math, and more than that, numbers are extraordinarily useful tools that allow us to understand and manipulate our natural world.

Back in the day, how did you learn that $5+6=11$? First, you defined the numbers. You started with 0, then 1, then 2, then 3, and so on, always adding 1 and classifying them into different deciles: 10, 100, 1000, etc. At its core is counting: the natural numbers $\mathbb{N}$ are also called the **counting** numbers. You usually started with 1 and worked up, then came back for the special and important case of 0. These numbers only exist as a concept on a piece of paper, and in your mind: you can put out 2 or 3 fingers to represent these ideas, but at the end of the day, counting is a human construct— this is what allows us to go from 2 fingers to 3 fingers, and so on. Mathematically, this counting operation is called the **successor operation**. We express it as a map $\text{succ} : \mathbb{N} \to \mathbb{N}$ with $\text{succ}(n) = n+1$. Addition is built up from repeated application of the successor operation ($m+n = \text{succ}^m(n)$) and multiplication is built up from repeated application of addition ($m \times n = \sum_{i=1}^{m} n$).

This logic is a bit circular: we expanded out the successor operation and used that to define the natural numbers by starting with 0 and applying the successor operation, but the successor operation is only defined in terms of the natural numbers. What we need is to construct a system that can function identically (isomorphically) to the natural numbers, based on the properties that we want it to satisfy: these properties are called the **Peano**

axioms. We'll have to define the set $\mathbb{N}$ in terms of *objects that we have at our disposal: sets*. At the same time, we'll define a successor operation and show that it **generates** $\mathbb{N}$, in that if we start with 0 (or 1) and repeatedly apply the successor operation, we'll eventually get the whole set. This will also implicitly define addition on this abstract set $\mathbb{N}$. After this, we'll show that $(\mathbb{N}, +)$ satisfies the axioms that we want it to, the Peano axioms, to consider it a valid natural number system. At this point we have a system $(\mathbb{N}, +)$ which functions exactly as we'd hope the natural numbers would, except it's built up rigorously from the ground up with mathematical rules that we know are consistent.

My two main points here: numbers didn't come for free, they weren't premade concepts, they were concepts that we, as humans, constructed. We had to construct the set of numbers as a whole. We already have a notion of 1, 2, 3, and so on, but defining operations on an abstract set of "letters" is hard— you need induction for that, and to define induction you need to have a system of counting already in place! The reason for this is logical consistency: I need to show that, for example, 3 + 4 always equals 7. Sure, that's obvious for 3+4, but what about 7362628383 + 36646282825? You need to define the addition operation, and to do that you need to define the successor operation, and to do that you need to understand what a number is from the ground up. It's not just enough to say "7362628383 means I have 7362628383 apples".

From a mathematical perspective, this is how we approach any object we wish to use, no matter how intuitive the object is: to start using other objects, we need to construct them explicitly in terms of objects that we've previously constructed. This applies to even the simplest object we can consider, the object that everyone with fingers can intuitively work with: the natural numbers $\mathbb{N}$. There are two stages to constructing a number system.

1. Identify properties you want the number system to satisfy. Any system with these properties will either be your number system, or it will generalize your number system.

2. Construct a system that satisfies the properties that you want. This system must be constructed in terms of objects that you've already constructed.

The integers are like the natural numbers, but they also have negative numbers. Back in 6th grade when you learned about negative numbers, they were probably extremely confusing: how can you have a negative of something? You can't hold -1 in your hand. This is really the first abstract mathematical concept that you have to grasp in your life: how can there be a -1 if I can't point to some amount of something and say "there are -1 of these objects there?"

To deal with this, you expanded your definition of numbers. You treated -1 as an abstract entity: treated on the same footing as a number, but not representable by an "amount" of an object you can hold. Instead, -1, and the rest of the negative numbers, were defined by their properties as an /textbfadditive inverse. Sure, it might have been a silly concept conceptually, but once you saw the utility that negative numbers had in solving arithmetic equations, you started to understand and appreciate it for what it was. Soon you likely stopped negative numbers as different than positive numbers and probably put them on an equal footing.

The same thing happened with the rational numbers. We defined multiplication through addition, then wanted an inverse. The inverse that we got was fractions, which proved to

have an immense amount of utility in solving algebraic equations. After you learned and got comfortable with fractions in fifth grade, you stopped thinking it was weird to have 1/12 of something, despite the fact that you can't count that you can't count to it or express it on your fingers.

We'll now build up the different number systems rigorously. Anything we construct must be mathematically consistent: because of this, we'll start with the only axiomatic theory we currently have, which is set theory. This means we will construct the natural numbers explicitly *from sets*: that will make sure that these objects indeed exist. We'll define a set of axioms, called the *Peano Axioms*, that we want the natural numbers to satisfy, and show that our system satisfies the Peano Axioms. At this point, we'll drop this granular view of the number system and refer to each number by its usual name that everyone is familiar with.

Interlude 3: Construction of the Natural Numbers

The natural numbers $\mathbb{N}$ are the set $\mathbb{N} := \{0, 1, 2, 3, ...\}$. The set is equipped with a unary operation called **successor**, denoted $n \mapsto \text{succ} : \mathbb{N} \to \mathbb{N}$, which effectively adds 1 to n, $\text{succ}(n) = n + 1$. The pair $(\mathbb{N}, \text{succ})$ is defined abstractly as a system which satisfies the **Peano axioms**:

1. Zero is a number.
2. Zero is not the successor of any number.
3. If $\text{succ}(a) = \text{succ}(b)$ for $a, b \in \mathbb{N}$, then $a = b$.
4. If a subset S of $\mathbb{N}$ contains 0 and, for each $n \in S$, S contains $\text{succ}(n)$, then $S = \mathbb{N}$.

Given a successor operator $\text{succ}(n)$, we define addition as the binary operation $+ : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$,

$$a + b := \text{succ}^b(a), \tag{53}$$

where succ^b denotes that you compose succ with itself b times (so $2 + 3$ is defined as $\text{succ}(\text{succ}(\text{succ}(2)))$. One then must show that this definition has the properties that we know and love: commutativity ($a+b = b+a$) and associativity ($(a+b)+c = a+(b+c)$).

In axiomatic set theory, it is not enough just to state axioms: you must construct a system that actually follows the axioms. The only way to construct such a system is with sets. We need to construct a system that represents the natural numbers; that is, it must be in bijection with what we typically think of as the natural numbers, and it must have a successor operation. We identify

$$0 := \emptyset, \tag{54}$$

and define our successor as follows,

$$\text{succ}(A) := A \cup \{A\}. \tag{55}$$

We then define $1 := \text{succ}(0) = \emptyset \cup \{\emptyset\} = \{\emptyset\}$, $2 := \text{succ}(1) = 1 \cup \{1\} = \{\emptyset, \{\emptyset\}\}$, and so on,

$$n := \text{succ}^n(0). \tag{56}$$

Note that this definition implies that $n = \{1, 2, 3, ..., n-1\}$. Expressing the $<$ ordering is also then trivial, and we say have

$$n < m \text{ iff } n \in m. \tag{57}$$

You can then show that the tuple $(\mathbb{N}, \text{succ})$ satisfies the Peano axioms, and that this system can be identified with the natural numbers!

The remaining operations we need to define on $\mathbb{N}$ are the binary operations of addition $+ : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ and multiplication $\cdot : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$. It is clear that we will define addition by repeatedly applying the successor operation, then multiplication by repeatedly applying the addition operation. To formally do this for arbitrary $n \in \mathbb{N}$, one needs more machinery that we will not discuss here, but if you are interested, check out Section **??** once it's complete (the recursion theorem).

We built up the natural numbers to include the number 0, but conventionally it often does not depending on the context. We will try to be specific in these notes. I prefer to denote $\mathbb{N} := \{1, 2, 3, ...\}$ and $\mathbb{N}_0 := \{0, 1, 2, 3, ...\} = \{0\} \cup \mathbb{N}$, so we will adopt this convention, but it is perfectly fine to use other conventions as long as you are consistent!

Now, we defined a successor and said that $\mathbb{N}$ was the set containing 0 and all its successors. Can we do this? We haven't really dealt with infinite sets before, but the notion that we can simply say $\mathbb{N}_0 = \{\text{succ}^n(\emptyset) : n \in \mathbb{N}_0\}$ is extremely circular. Before constructing $\mathbb{N}$, we **do not have an infinite object** in our naïve set theory: but we need an infinite object in order to construct $\mathbb{N}$ by **infinitely** applying the successor operation.

This topic will be taken up further in Section **??**, but the answer is not trivial and in fact lies in the Zermelo-Fraenkel axioms, specifically the **axiom of infinity**. A set S is called **inductive** if it contains the empty set $\emptyset$ and is **closed under the successor operation** (note the operation is defined on sets, not just numbers). The axiom of infinity states that there exists an inductive set, and we take $\mathbb{N}$ to be the intersection of all inductive sets, hence because of this axiom, $\mathbb{N}$ indeed exists.

This may get you thinking: properties of $\mathbb{N}$ (specifically the Peano axiom, and in particular axiom 4) led us to be able to define induction. Are there other types of induction that we may be able to do with different number systems? The answer is yes: there is a generalized form of induction called **transfinite induction** that we will discuss if we have time to talk about ordinal numbers.

Let's get back to the number systems. After constructing $\mathbb{N}$, the next number system that we want to construct is the integers, $\mathbb{Z}$. The integers extend the natural numbers by adding an additive inverse to the system. The simplest way to do this is to stack two copies of $\mathbb{N}$ on top of one another and add 0: that is, to take $\mathbb{Z} = \mathbb{N}_0 \cup \mathbb{N}'$, where $\mathbb{N}'$ is a copy of $\mathbb{N}$ (where $\mathbb{N}$ here starts at 1, as in the convention we are adopting). Here we consider the negative numbers to be elements of $\mathbb{N}'$, i.e. -1 is identified as $1'$. One then needs to define the addition operation on $\mathbb{Z}$, which now incorporates these negative numbers. Clearly, $+$ is

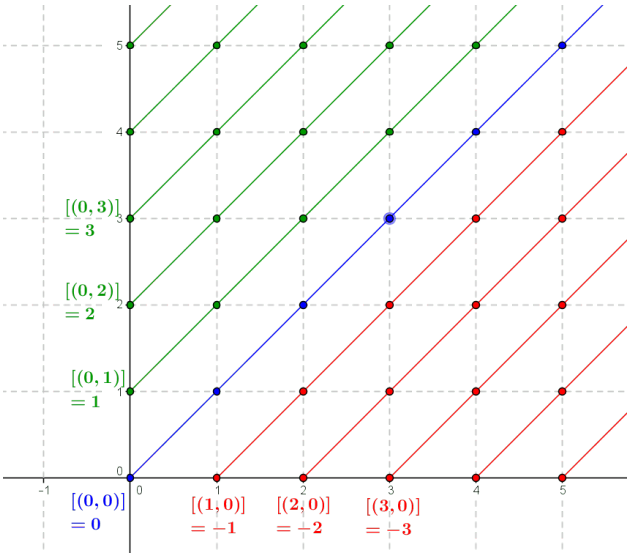

Figure 4: Construction of the integers from the natural numbers. The axes denote copies of $\mathbb{N}_0$: an integer is represented by an equivalence class of $\mathbb{N}_0 \times \mathbb{N}_0$ under the equivalence relation $(a, b) \sim (c, d)$ iff $a + d = b + c$.

defined on $\mathbb{N}_0 \times \mathbb{N}_0$ and $\mathbb{N}' \times \mathbb{N}'$, so one only needs to define the cross terms. This is where it gets difficult: the exact definition is not elegant, and $+$ as an operation is defined piecewise with a lot of different conditions. This means that it's hard to prove things with the system $(\mathbb{Z}, +)$, which is not a desirable property to have.

We'll instead provide an alternative construction of $\mathbb{Z}$ from $\mathbb{N}$ which is a bit more elegant and easier to prove things with. The idea is to view the integers as positive or negative intervals of natural numbers. The integer 2 will be represented by any pair $(n, m) \in \mathbb{N}_0^2$ with $n - m = 2$. Likewise, -2 will be represented by a pair with $n - m = -2$. We can use equivalence relations to impose that any two of these pairs are equal: this is shown in Figure 4. Let's do this construction formally.

Interlude 4: Construction of the set $\mathbb{Z}$ from $\mathbb{N}_0$

We will now construct the integers from $\mathbb{N}_0$, as depicted in Figure 4. As mentioned, this is done by viewing an integer $n \in \mathbb{Z}$ as a pair (a, b) of natural numbers with z being the size of the interval $a - b$. The subtlety here is that a and b are defined as natural numbers, and $-$ is not a well-defined operation on $\mathbb{N}_0$: for example, we want $1 - 2$ to equal -1, but -1 is in $\mathbb{Z} \setminus \mathbb{N}_0$, which has not yet been defined. We thus have to define $\mathbb{Z}$ from $\mathbb{N}_0$ only using the $+$ operation. We define an equivalence relation $\sim$ on $\mathbb{N}_0^2$ as,

$$(a, b) \sim (c, d) \text{ iff } a + d = b + c \tag{58}$$

Notice that *if* we had access to $-$, this would be $a - b = c - d$, which is exactly what we want. The integers $\mathbb{Z}$ are then defined as

$$\mathbb{Z} := \mathbb{N}_0 \times \mathbb{N}_0 / \sim . \tag{59}$$

The next step is to define the operations that we can put on $\mathbb{Z}$, namely addition, negation, and multiplication. We will also define an order $<$ on the $\mathbb{Z}$. For addition, we add intervals by just adding the top ends together and subtracting the bottom ends together: this gives us the definition of $+ : \mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}$ as,

$$[(a, b)] + [(c, d)] := [(a + c, b + d)]. \tag{60}$$

Negation is defined as simply swapping the order of the endpoints of each interval,

$$-[(a, b)] := [(b, a)], \tag{61}$$

where we can then define subtraction as $n - m := n + (-m)$, with $n, m \in \mathbb{Z}$. Next, we need to define multiplication of integers. We can figure out how multiplication is done by seeing how intervals would multiply together: we want $(a-b)(c-d) = ac-bc-ad+bd$, so the first component of $[(a, b)] \times [(c, d)]$ is $ac+bd$, and the second (negative) component is $ad + bc$, as follows:

$$[(a, b)] \times [(c, d)] := [(ac + bd, ad + bc)]. \tag{62}$$

Finally, we need to put an ordering on $\mathbb{Z}$. This is, of course, done by putting an ordering on the intervals that each one represents: that is, $[(a, b)] < [(c, d)]$ means that $a - b < c - d$, but we must rearrange these intervals to only use addition, hence

$$[(a, b)] < [(c, d)] \text{ iff } a + d < b + c. \tag{63}$$

The last thing that remains is to show these operations are **well-defined**. We defined all our operations in terms of *representatives* (a, b) of each equivalence class $[(a, b)]$. The equivalence class contains infinitely many elements, not just (a, b). So, we need to make sure that any representative that we choose also yields the same definition of these operations. For example, consider negation, which we defined as $-[(a, b)] = [(b, a)]$. We need to make sure that if we choose another representative, say (c, d), of the equivalence class $[(a, b)]$, then it has the same negation. Any $(c, d) \in [(a, b)]$ satisfies $c+b = d+a$. So, if we look at $-[(c, d)]$, we have $-[(c, d)] = [(d, c)]$. The question is thus: does $[(d, c)] = [(b, a)]$? Indeed it does, as $d + a = b + c$ implies $(d, c) \sim (b, a)$, hence $-[(c, d)] = [(d, c)] = [(b, a)] = -[(a, b)]$. This means that the negation operation is well-defined (defined independently of its representative). We need to show that each of the operations we defined are well-defined: I would suggest trying it out for addition to get the hang of it.

The next thing that we need to do is construct the rational numbers from the integers. Recall the rational numbers extend the integers by adding in multiplicative inverses for each element other than zero: we append $\mathbb{Z}$ with all possible fractions except for $1/0$. We will soon see that this is an example of a more general construction, called the **field of fractions**. A unital ring is a set equipped with two operations– addition and multiplication– that satisfies certain properties like distributivity, associativity, having an additive inverse $(-n)$, and having an identity element under addition (0) and multiplication (1). The integers are

a great example of a unital ring. Any unital ring can be extended to an object called a **field**, which has these properties, but additionally contains *multiplicative inverses* ($\frac{1}{n}$) for each non-zero element in the ring. The smallest such field that contains a ring R is called the field of fractions of R: $\mathbb{Q}$ is the field of fractions of $\mathbb{Z}$, and we'll see many more examples of this later. For now, let's move on to the construction of $\mathbb{Q}$ from $\mathbb{Z}$.

Interlude 5: Construction of the set $\mathbb{Q}$ from $\mathbb{Z}$

Recall that the rational numbers $\mathbb{Q}$ are of the form $\{\frac{a}{b} : a, b \in \mathbb{Z}, b \neq 0\}$. How do we construct this mathematically from the integers $\mathbb{Z}$? We could say that we have a bijection f between $\mathbb{Z} \times \mathbb{Z} \setminus \{0\}$ and $\mathbb{Q}$ given by

$$f : (a, b) \mapsto \frac{a}{b}. \tag{64}$$

However, this isn't quite right. Think for a second about what the problem here is: what is $f(2, 5)$, and what is $f(4, 10)$?

The problem with this bijection is that different objects in $\mathbb{Z}\times\mathbb{Z}\setminus\{0\}$ are **identified** with one another in $\mathbb{Q}$: we have $f(a, b) = f(ca, cb)$ for any $c \in \mathbb{Z}$ and $(a, b) \in \mathbb{Z} \setminus \{0\}$. This is the mathematical statement that says that we can reduce fractions: $1/2 = 2/4$ in $\mathbb{Q}$, for example. We will encounter this very often as we move through this course: $\mathbb{Z}\times\mathbb{Z}\setminus\{0\}$ "contains" the information we need about $\mathbb{Q}$, but it has too much information: we need to reduce the amount of elements in the set. This is expressed mathematically because the map f is surjective, but *not* injective. The simplest way to deal with this is to just take the simplest form representative in $\mathbb{Z}\times\mathbb{Z}\setminus\{0\}$ and throw the rest away, i.e. take $1/2$ and throw out $2/4, 4/8, ...$, but this isn't elegant and doesn't generalize well.

What we'll do now is to construct an equivalence relation on $\mathbb{Z}\times\mathbb{Z}\setminus\{0\}$ and quotient out by the equivalence classes. Two fractions $\frac{a}{b}$ and $\frac{c}{d}$ are equivalent when $ad = bc$, so this is what we'll quotient out by: we define the relation $\sim$ on $\mathbb{Z} \times \mathbb{Z} \setminus \{0\}$ by,

$$\frac{a}{b} \sim \frac{c}{d} \text{ iff } ad = bc. \tag{65}$$

This is an equivalence relation. Reflexivity and symmetry are immediate, and transitivity is because if $\frac{a}{b} \sim \frac{c}{d}$ and $\frac{c}{d} = \frac{g}{h}$, then $ad = bc$ and $ch = dg$, so $(ad)(ch) = (bc)(dg) \implies (ah)(cd) = (bg)(cd) \implies ah = bg$, hence $\frac{a}{b} \sim \frac{g}{h}$. Now, we identify $\mathbb{Q}$ as

$$\mathbb{Q} := (\mathbb{Z} \times \mathbb{Z} \setminus \{0\})/\sim \tag{66}$$

We'll often express a fraction as $[\frac{a}{b}]$, which represents its coset under $\sim$, i.e. $[\frac{a}{b}] = \{q \in \mathbb{Z} \times \mathbb{Z} \setminus \{0\} : q \sim \frac{a}{b}\}$. We see that quotienting by $\sim$ exactly replicates the structure of $\mathbb{Q}$ that we wanted, which is the ability to reduce fractions and multiply by a common factor! Modding out by equivalence relations is exactly is how we will construct sets from other sets by saying that given elements are equal (*identifying* these different elements).

The last thing to discuss now (we'll discuss this construction at length later when we talk about algebra) is how we add elements in $\mathbb{Q}$. We define addition as:

$$[(a,b)] + [(c,d)] := [(ad+bc, bd)] \tag{67}$$

Intuitively, this is just fraction addition $\frac{a}{b} + \frac{c}{d} = \frac{ad+bc}{bd}$. The hard part of this definition is showing that it's **well-defined**: that $(a,b)+(c,d)$ is always in the same coset, assuming (a,b) are arbitrary elements of the coset $[(a,b)]$ (and likewise for (c,d)). So, we let $(a,b) \sim (g,h) \in [(a,b)]$ and $(c,d) \sim (i,j) \in [(i,j)]$, and we need to show that $(ad+bc, bd) \sim (gj+ih, hj)$. Using that $ah = bg$ and $cj = di$, we have:

$$(ad+bc)(hj) = adhj + bchj = (dj)(bg) + (bh)(di) = (gj+ih)(bd) \tag{68}$$

Make sure you can follow this, then do the equivalent proof for multiplication!

1.6 What is a number? $\mathbb{R}$

The hardest number system to define is the real numbers. This is because $\mathbb{R}$ shares all the same algebraic properties as $\mathbb{Q}$: it is an ordered field in the same way that $\mathbb{Q}$ is. In buildilng up the previous number systems $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q}$, we defined each number system sequentially by imposing some algebraic property. For $\mathbb{N} \subset \mathbb{Z}$, we constructed $\mathbb{Z}$ to be a group under $+$ by adding in additive inverses, and for $\mathbb{Z} \subset \mathbb{Q}$, we constructed $(\mathbb{Q}, +, \cdot)$ to be a field by adding in multiplicative inverses. There is no simple leap from $\mathbb{Q}$ to $\mathbb{R}$.

Instead, $\mathbb{R}$ extends $\mathbb{Q}$ by adding in *topological properties*. These properties are harder to define, but you'll pick up on what they mean intuitively. The property that $\mathbb{R}$ adds to $\mathbb{Q}$ is called **completeness**: it essentially "fills in" the holes in the real number line that $\mathbb{Q}$ doesn't touch. The usual formulation of the completeness axiom specifies this by stating that every subset which is bounded above has a "least upper bound": an upper bound which is smaller than all other upper bounds. This is the notion of the **supremum** and the **infimum** of a subset of the real numbers.

Definition 5: Bounded

Let $A \subseteq \mathbb{R}$ be a non-empty set. We say A is **bounded above** if there exists $M \in \mathbb{R}$ such that $a \leq M$ for each $a \in A$. Likewise, A is **bounded below** if there exists $M \in \mathbb{R}$ with $M \leq a$ for each $a \in A$.

Definition 6: Supremum, Infimum

Let $A \subseteq \mathbb{R}$ be a non-empty subset. The **supremum** of A, denoted $\sup A$, is its *least upper bound*. That is, $\sup A \geq a$ for each $a \in A$, and for any other upper bound $M \in \mathbb{R}$ (so any M such that $M \geq a$ for each $a \in A$), $\sup A \leq M$. The **infimum** of A, denoted $\inf A$, is its *greatest lower bound*, defined analagously.

The supremum and infimum generalize the notion of min and max to infinite subset of $\mathbb{R}$. On finite sets, the supremum agrees with the maximum, and the infimum agrees with the minimum. For example, for $A = \{4, 21, 0, -2, -62, 315\}$, we have $\max A = \sup A = 315$ and $\min A = \inf A = -62$.

Exercise 3

Let $A \subseteq \mathbb{R}$ be a non-zero, finite subset. Denote $A = (a_i)_{i=1}^N$. Prove that

$$\sup A = \max_{i=1}^{N} a_i \qquad \inf A = \min_{i=1}^{N} a_i. \tag{69}$$

Where the supremum and infimum come in handy is when we deal with infinite sets and intervals. Consider the intervals $A = (0, 2)$, $B = [0, 2)$, and $C = [0, 2]$. The maximum and minimum of these sets are only defined when the corresponding endpoint is closed: $\min A, \max A$, and $\max B$ are all undefined, while $\min B = 0, \min C = 0$, and $\max C = 2$.

However, the supremum and infimum of such sets *are defined*, and correspond exactly with what we'd expect the min / max to be,

$$\inf A = 0 \qquad \sup A = 2 \qquad \sup B = 2. \tag{70}$$

The supremum and infimum can also be infinite, for example $\sup[0, \infty) = \infty$. We will be using these two operations extensively any time we discuss the real numbers. With these definitions, we can now state the **completeness axiom** of $\mathbb{R}$. We'll also state an important corollary of the completeness axiom.

Definition 7: Completeness Axiom of $\mathbb{R}$

Every non-empty subset $S \subseteq \mathbb{R}$ which is bounded above has a least upper bound: $\sup S$ exists and is finite. Likewise, every non-empty subset $S \subseteq \mathbb{R}$ which is bounded below has an infimum, and $\inf S$ is finite.

Corollary 1: Denseness of $\mathbb{Q}$ in $\mathbb{R}$

Let $a, b \in \mathbb{R}$ with $a < b$. Then, there exists $q \in \mathbb{Q}$ with

$$a < q < b. \tag{71}$$

Let's illustrate why we need the completeness axiom, and how $\mathbb{R}$ is different than $\mathbb{Q}$. Consider the set $S := (-\infty, \sqrt{2}) \cap \mathbb{Q} \subset \mathbb{Q}$ in both $\mathbb{Q}$ and $\mathbb{R}$. The supremum of this set is $\sqrt{2} \notin \mathbb{Q}$, despite the $S \subset \mathbb{Q}$. If we were to try to find a rational supremum for S, we would fail because of the density of $\mathbb{Q}$. Suppose that $q \in \mathbb{Q}$ is the supremum of S. The important point is that $q \neq \sqrt{2}$, so either $q \in \sqrt{2}$ or $\sqrt{2} < q$. In the first case, then q is not even an upper bound of S, because $q < \sqrt{2}$ implies that $\exists q' \in \mathbb{R}$ such that $q < q' < \sqrt{2}$ (by the density of $\mathbb{Q}$), and $q' \in S$ by definition with $q < q'$. In the second case, then q is not the **least** upper bound, because again $\sqrt{2} < q$ implies there is $q'' \in \mathbb{Q}$ with $\sqrt{2} < q'' < q$, and q'' is a smaller upper bound for S than q. Hence, the supremum of S **cannot be a rational number, despite** S **being a subset of the rational numbers**. The supremum of S is instead $\sqrt{2}$. Clearly by definition $\sqrt{2}$ is an upper bound, since S is defined to be all rational numbers less than $\sqrt{2}$. Note that $\sqrt{2}$ must be the least upper bound, because for any $M < \sqrt{2}$, $M \in S$ and by the density of $\mathbb{Q}$, there exists $q \in \mathbb{Q}$ with $M < q < \sqrt{2}$, i.e. $q \in S$ and $M < q$ means that M cannot be an upper bound. Hence $\sup S = \sqrt{2}$, even though S was composed completely of rational numbers.

Let's get back to the problem of constructing $\mathbb{R}$ from $\mathbb{Q}$. There are many different equivalent constructions, and they are all quite complicated. We'll see a different construction soon when we do a bit of real analysis that I prefer over this one, but for now we'll work with a construction called a **Dedekind cut**. The idea behind Dedekind cuts is that each real number $r \in \mathbb{R}$ corresponds uniquely to a half-infinite $\mathbb{Q}$-valued interval

$$\alpha_r := (-\infty, r) \cap \mathbb{Q} = \{q \in \mathbb{Q} : q < r\} \subseteq \mathbb{Q}. \tag{72}$$

Each set of this form uniquely corresponds to a real number: if $r \neq r'$ (WLOG take $r < r'$), then the density of $\mathbb{Q}$ implies $\exists q \in \mathbb{Q}$ such that $r < q < r'$, and we see that $q \in \alpha_{r'} \setminus \alpha_r$.

Now, this logic is of course circular right now. We defined α_r using a real number r, but we have not yet constructed the real numbers, so we have no r to use for this definition. Instead, we'll formally construct Dedekind cuts to satisfy a few given postulates, then show that we have the desired set. To do this, we need to abstract away the properties that we want Dedekind cuts to have. We want them to be non-trivial, i.e. not $\emptyset$ and not $\mathbb{Q}$. We also want them to be downard closed, so they are of the form $(-\infty, r) \cap \mathbb{Q}$ and stretch all the way down to $-\infty$. Finally, each of the Dedekind cuts we looked at had an open interval on its top end, i.e. it cannot have a maximum element. Let's now formalize this discussion.

Interlude 6: Construction of $\mathbb{R}$ from $\mathbb{Q}$: Dedekind Cuts

A subset $\alpha \subseteq \mathbb{Q}$ is called a **Dedekind cut** if

1. α is non-trivial: $S \neq \emptyset$ and $S \neq \mathbb{Q}$.
2. α is *downward closed*: if $r \in \alpha$ and $s \in \mathbb{Q}$ with $s < r$, then $s \in \alpha$.
3. α has *no largest element*.

Any set with these properties is intuitively of the form $(-\infty, r) \cap \mathbb{Q}$. Because we argued that such sets were in bijection with the real numbers, we use them to **define** the real number line. The real number system $\mathbb{R}$ is hence **defined** as the set of all Dedekind cuts,

$$\mathbb{R} := \{\alpha \subseteq \mathbb{Q} : \alpha \text{ is a Dedekind cut}\}, \tag{73}$$

where intuitively, the "value" of the Dedekind cut α is the number it extends to on its upper bound. For example, the set $\alpha_{1/2} = (-\infty, 1/2) \cap \mathbb{Q}$ is a Dedekind cut which represents the number $1/2 \in \mathbb{Q}$: when we view $1/2$ as a real number, this is actually what we're using under the hood. Said formally, there is an **embedding** of $\mathbb{Q} \hookrightarrow \mathbb{R}$ given by $q \mapsto \alpha_q$, where $\alpha_q := (-\infty, q) \cap \mathbb{Q}$ is the corresponding Dedekind cut. Thus we can consider the rational number $q \in \mathbb{Q}$ as either the previous construction we did (cosets of $\mathbb{Z} \times \mathbb{Z} \setminus \{0\}$ under $\sim$), or as a subset of $\mathbb{Q}$ of the specific form $(-\infty, q) \cap \mathbb{Q}$.

We have now defined the set of real numbers $\mathbb{R}$, but we still have not defined an order on $\mathbb{R}$ nor operations. We start with $<$: since a Dedekind cut is downward closed, we can define

$$\alpha < \beta \text{ iff } \alpha \subseteq \beta \tag{74}$$

(if this is confusing, use the usual parameterization of $\alpha_r = (-\infty, r) \cap \mathbb{Q}$).

Addition is defined element-wise for each pair of Dedekind cuts,

$$\alpha + \beta := \{a + b \in \mathbb{Q} : a \in \alpha, b \in \beta\}. \tag{75}$$

One must prove that $\alpha + \beta$ is indeed a Dedekind cut before proving the other properties of addition (associativity, commutativity, identity and additive inverse). Multiplication is trickier, and we will not define it here. The reason is that negative numbers do not act nicely with the element-wise definition of multiplication. Suppose that one defined $\alpha \cdot \beta := \{a \cdot b \in \mathbb{Q} : a \in \alpha, b \in \beta\}$. What would happen if we multiplied a Dedekind cut α by -1? This would **not** be a Dedekind cut under element-wise multiplication, since

-1 would flip the Dedekind cut α to be *upward-closed*. So, multiplication is more subtle: we will not construct it explicitly here, but assume it works and behaves as expected.

Equipped with $(+, \cdot, <)$, $\mathbb{R}$ is now an ordered field. There are a few final important things to prove that you should try if you're interested. The first is that the canonical map $\mathbb{Q} \hookrightarrow \mathbb{R}, q \mapsto \alpha_q$ respects the ordered field structure of $\mathbb{Q}$ and $\mathbb{R}$: one needs to prove this map is a **field homomorphism** (we already know it is injective) and preserves the ordering $\leq$. A field homomorphism is a map which respects the field structure of these two spaces: that is, $\alpha_{q+q'} = \alpha_q + \alpha_{q'}$ and $\alpha_{q \cdot q'} = \alpha_q \cdot \alpha_{q'}$. The last important thing to prove is that $(\mathbb{R}, +, \cdot)$ is **complete**: that every bounded subset has a supremum. We will not show this, because the proof is very involved. But, this is perhaps the most important thing to prove from this construction, as it's the motivation for constructing $\mathbb{R}$! Later in Chapter **??** we will discuss another way to view the concept of completeness and revisit some of this discussion.