PHL349
Set Theory
by Brandon Papandrea

The following text is based on lecture notes written during the Winter 2026 offering of PHL349 - Set Theory at the University of Toronto, and are in large part based on the textbook *Elements of Set Theory* by Herbet Enderton, but also include notes based on the books *Conceptions of Sets and the Foundations of Mathematics* by Luca Incurvati, and *The Joy of Sets* by Kieth Delvin. The inention for these notes is to be a more polished and complete version of my own handwritten notes I take during class, and thus are not intended to be a primary source for learning set theory.

Contents

1 Lecture 1: Two Conceptions of Sets 1

1.1 Why Should I Care About Sets? . . . 1
1.2 What is a Set? . . . 1
1.2.1 The Logicist Conception of Sets . . . 1
1.2.2 The Generative Conception of Sets . . . 4
1.2.3 Which is Better? . . . 5
1.2.4 The Arithmetical Conception . . . 6

2 Lecture 2: The ABCs of Set Theory 7

2.1 Notations and First Axioms . . . 7
2.2 The Algebra of Sets . . . 8
2.3 Functions and Relations . . . 10
2.3.1 Ordered Pairs . . . 10
2.3.2 Cartesian Products, Binary Relations, and Generalizations . . . 10
2.3.3 Functions . . . 12
2.4 The Axiom of Choice (AC) . . . 12
2.4.1 First Form . . . 12
2.4.2 Second Form . . . 13

3 Lecture 3: Constructing the Naturals 14

3.1 The Existence of the Naturals . . . 14
3.1.1 Simple Infinite Systems and the Axiom of Infinity . . . 14
3.1.2 The Dedekind-Peano Postulates . . . 15
3.2 The Elementary Theory of Arithmetic . . . 19
3.2.1 Addition and Multiplication . . . 19
3.2.2 Ordering the Naturals . . . 20

4 Lecture 4: Number Systems 22

4.1 The Integers . . . 22

4.2 The Rationals 25
4.3 The Reals 27

5 The Size of Sets 29

5.1 Equinumerosity 29
5.2 Finite Sets 31
5.3 Cardinal Arithmetic 32
5.4 Ordering Cardinal Numbers 33

1

Lecture 1: Two Conceptions of Sets

In the opening lecture of this course, we seek to understand what sets are through a philosophical lens by considering two different ways of conceptualizing them. Each have their pros and cons.

1.1 WHY SHOULD I CARE ABOUT SETS?

Any discussion about sets, especially at the level of formal set theory, will eventually lead to someone asking "what on Earth is the point of all this?" This is reasonable to ask considering how most math students, even those seeking to do research in the subject, will likely never have to deal with sets in this manner.

The reason why we care about sets is, in short, because sets have a maximum unifying power in mathematics. There are two facts that make this true:

1. Every concept in math may be defined in terms of sets, and all mathematical statements can be expressed using the language of set theory (a first-order language that has a single binary relation, membership).
2. All of mainstream math can be developed within the theory of sets, which is known as **Zermelo-Fraenkel with Choice**, or just **ZFC**.

The first fact follows from the power of the first-order language, which we call $\mathcal{L}_\varepsilon$, and that such first-order languages are sufficient in expressing mathematical reasoning. The second fact comes from how ZFC was specially constructed to allow us to formalize mathematics. Mathematicians and philosophers like Dedekind, Cantor, Frege, Russell, Whitehead, and finally, Zermelo & Fraenkel, worked over decades to find this natural theory in the late 19th and early 20th centuries. The fact that ZFC was chosen and not some other theory has a complicated history that we won't go into.

1.2 WHAT IS A SET?

The question of the day is the most basic one can ask in set theory: what exactly *is* a set? Answering this is not at all easy. We will discuss two schools of thought. In one, we say sets are defined, while in the other, we say they are constructed.

1.2.1 THE LOGICIST CONCEPTION OF SETS

Georg Cantor defined sets as a "totality of definite elements that can be combined into a whole by a law." What did he mean by "law?" To some, like Frege and Crispin Wright, a law is

a concept, or in other words, like a predicate in first-order logic. This forms the basis of the **logicist** conception of a set.

The key idea in this school of thought is as follows: given any predicate/concept, there exist a set containing only those things to which the predicate applies. This set is called the **extension of the predicate**. In this way, set theory is just a part of logic, like all other parts of math.

This conception of sets is based around two fundamental principles that each answer two important questions:

1. **Existence**: What sets are there?
2. **Identity**: How can we distinguish between two sets?

We can formalize these questions, and hence this conception of sets, in logical terms. Let $\mathcal{K}$ be a suitably powerful language, whose variables range over all sets and objects that are not sets (called individuals). Let S mean 'is a set,' and $\in$ mean 'is a member of.' The logicist principle of set existence can then be expressed as follows: If ϕ is a formula of $\mathcal{K}$, then there should be a set of just the things to which the predicate applies. In logic terms, if $\phi(x)$ is some well-formed formula of $\mathcal{K}$, we write

$$\exists y[Sy \wedge \forall x(x \in y \iff \phi(x))]$$

in the logicist theory, this is an axiom. In broader terms, any instance of the *comprehension schema* show up in the logicist theory as *existence postulates*.

The second principle is an identity principle. Think of how one would differentiate between two sets of, say, pencil crayons. If you can show that one set contains a coloured pencil not found in the other set, then we know that the sets are different. From the logicist viewpoint, this idea - differences in membership - is the only way to distinguish between sets:

$$\forall x \forall y[(Sx \wedge Sy \wedge x \neq) \implies \exists z((z \in x \wedge z \notin y) \vee (z \notin x \wedge z \in y))]$$

We can also write this another way: if two sets have exactly the same elements, they're the same set:

$$\forall x \forall y[(Sx \wedge Sy) \implies (\forall z(z \in x \iff z \in y) \implies x = y)]$$

This idea is called *extensionality*.

By combining the comprehension schema and the extensionality schema, we get a theory commonly referred to as **naïve set theory** (you'll see in a second why we use the word naïve). Some of its axioms are as follows:

1. $\exists y[Sy \wedge \forall x(x \in y \iff x \neq x)]$ (There is an empty set)
2. $\forall z \forall w \exists y[Sy \wedge \forall x(x \in y \iff (x = z \vee x = w))]$ (There is a pair set)
3. $\forall z \exists y[Sy \wedge \forall x(x \in y \iff \exists w(w \in z \wedge x \in w))]$ (There is a union set)
4. $\exists y[Sy \wedge \forall x(x \in y \iff (Sx \wedge x = x))]$ (There is a universal set)

Russell's Paradox

Frege didn't state his logicist theory like this. Instead he used a second-order language and made two logical assumptions:

1. $\exists X \forall y(Xy \iff \varphi(y))$ for all second-order well-formed formulas φ (comprehension schema)
2. $X \equiv Y \iff \forall z(Xz \iff Yz)$ (extensionality schema)

But, Frege made a very small yet fatal error. He made two additional assumptions about the existence of concept extensions. He assumed that every concept has an extension. More importantly, he assumed that distinct concepts have distinct extensions. This is known as **Law V**. comprehension and extensionality for sets then comes from defining a set as just the objects that are concept extensions, and membership in terms of extensions and predication: If we let Qxy mean 'x is an extension of Y,' then

1. $Sx := \exists Y[QxY]$ (an object is a set if it extends some concept)
2. $x \in y := \exists Y[QyY \wedge Yx]$ (an element of a set is an object that satisfies the corresponding concept)

The effect of these assumptions is that there are as many objects (sets), as there are concepts (predicates). But this does not fall in line with Cantor's theory that the power set of a set is larger than the set itself.

This leads us to the famous **Russell's Paradox**, which shows that naïve set theory is inconsistent. We consider the well-formed formula

$$Uy \wedge \forall x[Rxy \iff (Ux \wedge \neg Rxx)]$$

We can break this up into two parts at the outside $\wedge$ symbol:

$$Uy \quad \text{and} \quad \forall x[Rxy \iff (Ux \wedge \neg Rxx)]$$

So Uy is true. We instantiate the right side by setting all x's to be y's. Giving us

$$Ryy \iff (Uy \wedge \neg Ryy)$$

as Uy is true, we just get that

$$Ryy \iff \neg Ryy$$

a clear contradiction. This means that the negation of this entire formula is true:

$$\neg \exists y[Sy \wedge \forall x[x \in y \iff (Sx \wedge x \notin x)]]$$

but notice that this is an instance of the comprehension schema. We have that this instance of the scheme is both true and false, hence the theory is inconsistent. This idea was first shown to Frege via a letter written by Bertrand Russell in 1902.

There are a handful of ways people have thought of to fix this problem. The first way involves modifying our comprehension schema to make it more predicative. One may argue that, since we are defining a set using a quantifier that runs over all sets, which would include the set we're defining, we are using circular logic. To remedy this, we can either ban the use of quantifiers in our scheme, or use some sort of level system: variables are put into levels $0, 1, 2, \ldots, n, \ldots$ Then a set of level n may only be defined using variables of levels below n. This is quite complicated, and leads to more complexity than its worth, at least for some.

Another way is to look back at Frege's assumption and modify it: instead of saying all concepts have extensions, we say that only "safe" ones do. What is "safe?" That takes some work, and many of done work in this area. Some, however, have argued that the idea that certain concepts have no extensions, called the **limitation of size** is unnatural. In general, the problem of distinguishing between concepts that have extensions, and those that do not, is called the **Bad Company Problem**.

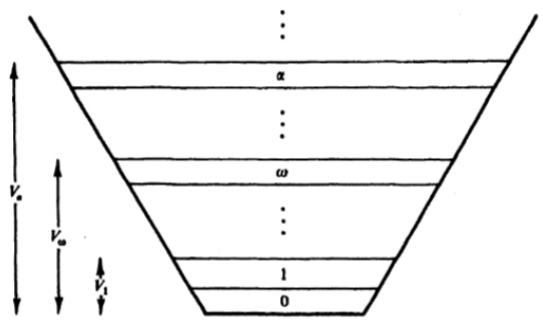
1.2.2 THE GENERATIVE CONCEPTION OF SETS

The second viewpoint we will discuss is to think of sets as not defined, but rather, *generated* in some iterative process. We can talk about this process using informal language, like 'stage,' 'is formed at,' 'earlier than,' etc.

We start with individuals, things that are not sets, and at Stage 0, form all possible collections of these individuals. If none exist, we only form the null set.

At Stage 1, we form all collections of individuals and sets formed at Stage 0. At Stage 1, we form all collections of individuals and sets formed at Stage 1. We keep going like this, and eventually reach Stage ω, the stage immediately stages 0,1,2, and so on. We can still keep going, performing transfinite interation, with stages $\omega + 1, \omega + 2, \ldots, 2\omega, \ldots, 3\omega, \ldots, \omega^2, \ldots, \omega^\omega, \ldots$.

One can visualize this as an infinite large cone. At the base we have the collections of just individuals; Stage 0. Each level of the cone is a stage, which is built on top of the stages built previously.



Under the generative conception, every set is repeatedly formed at each stage. More importantly, no set is an element of itself, as a set is formed at some earliest stage, and its elements must have been formed earlier than that. Thus, there is no universal set. Using a similar argument, there are no sets x, y such that $x \in y$ and $y \in x$. Thus, the set of all sets that do not contain themselves is just the set of all sets, which we know cannot exist. So, under the generative conception, Russell's Paradox is avoided!

The Theory of Stages

Because we have introduced the idea of stages, we need to formalize the theory of stages in logic. Our language $\mathcal{S}$ will have two types of variables: $x, y, z, w, \ldots$, which range over sets, while $r, s, t, \ldots$ will range over stages. We also include two new relational symbols: E for 'is earlier than,' and F for 'is formed at.' We can now state some axioms:

1. $\forall s \neg Ess$ (E is anti-reflexive)

2. $\forall r \forall s \forall t[(Ers \wedge Est) \implies Ert]$ (E is transitive)
3. $\forall s \forall t[Est \vee Ets \vee s = t]$ (E is connected)
4. $\exists s \forall t[s \neq t \implies Est]$ (There's an earliest stage)
5. $\forall s \exists t[Est \wedge \forall r(Esr \implies (Etr \vee t = r))]$ (Each stage has a unique successor stage)
6. $\exists s[\exists t Ets \wedge \forall t(Ets \implies \exists r(Etr \wedge Ers))]$ (There's a limit stage different from the first stage, but with no direct predecessor)
7. $\forall x \exists s[Fxs \wedge \forall t(Fxt \implies t = s)]$ (Sets form at some unique stage)
8. $\forall x \forall y \forall s \forall t[(y \in x \wedge Fxs \wedge Fyt) \implies Ets]$ (All elements of a set form before the set)
9. $\forall x \forall s \forall t[(Fxs \wedge Ets) \implies \exists y \exists r(y \in x \wedge Fyr \wedge (t = r \vee Etr))]$ (Sets form immediately after their elements are)

That ninth axiom is important, as it imitates the comprehension schema: at every stage all definable sets of previously formed elements are formed. How are they defined? Using the language $\mathcal{S}$. Formally, if $\chi(x)$ is an $\mathcal{S}$ formula that has no free occurrence of y, then

$$\forall s \exists y \forall x[x \in y \iff (\chi(x) \wedge \exists t(Ets \wedge Fxt))]$$

is an axiom of stage theory, called the **specification axiom**.

We also need to take care of one other thing: induction. Since we introduced stages in an inductive way, there better be a formal way of defining induction on stages, like there is a way of formally defining induction on numbers. We omit this here because it's not entirely useful for our purposes, but the work has been done and can be found.

The long and short of our discussion on stage theory is that it has a surprising amount of power. In fact, stage theory can prove the axioms of the empty set, pairing, union (in a general form), power-sets, and the axioms of infinity and separation (which says that for a set x there is a subset of elements of x satisfying a formula). This is almost enough to prove all of ZFC. However, there are some things stage theory can't prove that are required for ZFC. For example, the replacement schema, which says that the image of a set under a function is also a set. The axiom of extensionality is taken to be analytic, in the sense that it is an essential property of the notion of a set. Most importantly, the axiom of choice cannot be derived. Even still, stage theory can derive almost all of set theory, so it is quite powerful.

1.2.3 WHICH IS BETTER?

Is one of these conceptions better than the other, or is there a compramise we have to make? Some have argued that neither of these viewpoints is sufficient in justifying ZFC, and so we will have to make choices about which viewpoint to use depending on circumstances. How do we decide? That's a question outside the scope of the course, but the curious reader can find answers. In these notes we will spend most of our time working with the iterative conception of sets.

A particularly interesting concept that arises in both of these conceptions, more so in the logicist approach because of the limitation of size, is the idea that the set-theoretic universe in **ineffable**. We can talk about sets, and we know there is some set of all sets, but can we say anything about the entire universe we are working in. More concretely, consider the cone from

the generative conception. We can talk about individual sets in the cone or levels of the cone, but can we talk about the cone itself using the language of set theory? This idea actually has some applications to theology, since God acts similarly to this set-theoretic universe. Many have argued that, while God cannot be described using language, language can be used to approach God. There is an analagouse idea in set theory: while it may not be possible to describe our universe using the language of set theory, we can approach it. There is work to be done in this direction.

1.2.4 THE ARITHMETICAL CONCEPTION

There is one more notion of sets that is worth mentioning. In the **arithmetic conception** of sets, numbers form the basic mathematical objects. In this way, ZFC can be formed by combining two things: the theory behind finite collections (called Peano Arithmetic) and some notion of infinity. It is a well-known piece of folklore in math that Peano Arithmetic is equivalent to ZFC with the axiom of infinity negated.

2

Lecture 2: The ABCs of Set Theory

In this lecture, we introduce the basic tools of set theory, including the relevant conventions, terms, axioms, and set theoretic descriptions of relations and functions. We also describe an object called the Boolean algebra of operations on sets, and discuss a nice result about it: the Stone reprsentation theorem.

2.1 NOTATIONS AND FIRST AXIOMS

We are working in the first order language $\mathcal{L}_\in$ with identity, so $\in, =$ are relational symbols.

For sets, we use variables $x, y, z, \ldots$, while for classes, collections with objects as members, use variables A, B, C. Classes always appear to the right of membership ($x \in A$), and never to the left. Sets that are collections are objects and thus can be members of a class or set ($x \in y$).

Our first axiom is the axiom of **extensionality**, which essentially says that if two classes contain the same objects, then they are the same class:

$$\forall A \forall B[\forall x(x \in A \iff x \in B) \implies A = B]$$

This is an important axiom that allows us to define many well known sets. One of which is the **empty set**. The axiom defining it is

$$\exists B \forall x \; x \notin B$$

By extensionality, if we have two sets satisfying this axiom, then they contain the same members and thus are the same set. This is the empty set, denoted by $\varnothing$.

Remark. *The idea of an empty set is somewhat controversial, especially amongst ontologists. There does exist a theory without empty sets, ZFC+, that ZFC can be imposed into.*

Another set we may define is a class containing a pair of sets, called a **pair**:

$$\forall u \forall v \exists B \forall x[x \in B \iff (x = u \vee x = v)]$$

Again, by extensionality, this is unique, and we denote it as $\{u, v\}$. Note that if $u = v$, this gives us the singleton set $\{u\}$.

We may also wish to create sets which takes elements from two different sets and contains their unique elements. This forms the basis for unions. We state the **union axiom** in two forms. First the weak form:

$$\forall u \forall v \exists B \forall x [x \in B \iff (x \in u \vee x \in v)]$$

and by extensionality we may denote this as $x \cup y$.

We can define a partial relation on sets $\subseteq$, which we define as

$$x \subseteq a := \forall t (t \in x \implies t \in a)$$

in doing so, we can define the **power set axiom** as

$$\forall a \exists B \forall x [x \in B \iff x \subseteq a]$$

Extensionality means we can denote this set, the **power set** of a, as $\mathcal{P}a$. This is a much greater beast compared to the other sets we've defined. It is much larger and has no great way of intuitively being described.

To represent classes (and thus sets), our basic method is **abstraction notation**. For a formula $\phi(x) \in \mathcal{L}_{\in}$, we write

$$\{x | \phi(x)\}$$

as the class (collection) of sets that $\phi(x)$ is true of.

The **separation/subset** axiom allows us to separate elements in a set to creaete a new set:

$$\forall \vec{t} \forall a \exists B \forall x [x \in B \iff (x \in a \wedge \phi(x))]$$

where $\vec{t}$ are the parameters of $\phi(x)$. B is thus the set of those objects in a that $\phi(x)$ is true of. Applying this schema and extensionality with $\phi(x)$ being the formula defining another set b, or $\neg\phi(x)$, we get

$$a \cap b \quad \text{and} \quad a - b$$

We can use separation to get a more general notion of intersection as well:

$$\forall a [a \neq \varnothing \implies \exists B \forall x (x \in B \iff \forall y (y \in a \implies x \in y))]$$

This set B, the set of those objects that are in every object in the class a, is unique and we can write it as $\cap a$. We can also get an analogue general notion of union as well:

$$\forall a \exists B \forall x [x \in B \iff \exists y (y \in a \wedge x \in y)]$$

and we write B, the set of those objects that are in some object in a, as $\cup a$. Applying this axiom to the pairing axiom gives the weaker form of the union operation.

2.2 THE ALGEBRA OF SETS

We now have an algebraic structure $\langle V, \varnothing, \subseteq, \cap, \cup, - \rangle$, containing the universal class V, a minimal element $\varnothing$, a relation $\subseteq$, and operations $\cap, \cup, -$. Interestingly, this structure shares similarities with many well known mathematical and logical structures.

Given any classes A, B, C, the following may be proven:

– Commutativity

$$A \cup B = B \cup A \quad \text{and} \quad A \cap B = B \cap A$$

– Associativity

$$A \cup (B \cup C) = (A \cup B) \cup C$$
$$A \cap (B \cap C) = (A \cap B) \cap C$$

– Distributivity

$$A \cap (B \cap C) = (A \cap B) \cup (A \cap C)$$
$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

– De Morgan's Laws

$$C - (A \cup B) = (C - A) \cap (C - B)$$
$$C - (A \cap B) = (C - A) \cup (C - B)$$

– Empty Set Laws

$$A \cup \varnothing = A \quad \text{and} \quad A \cap \varnothing = \varnothing$$
$$A \cap (C - A) = \varnothing$$

– Montanacity and Anti-Montanacity Laws: If $A \subseteq B$,

$$A \cup C \subseteq B \cup C$$
$$A \cap C \subseteq B \cap C$$
$$\cup C \subseteq \cup B$$
$$C - B \subseteq C - A$$
$$A \neq \varnothing \implies \cap B \subseteq \cap A$$

– Distributivity (genearlized to arbitrary union and intersection)

$$B \neq \varnothing \implies A \cup (\cap B) = \cap\{A \cup x | x \in B\}$$
$$A \cap (\cup B) = \cup\{A \cap x | x \in B\}$$

– De Morgan's Laws (genearlized to arbitrary union and intersection)

$$C - \cup A = \cap\{C - x | x \in A\}$$
$$C - \cap A = \cup\{C - x | x \in A\}$$

A more interesting case is when these operations are restricted to some subset of V. Take for instance $\mathcal{P}A$ for a non-empty set A. Then we get a structure

$$\langle \mathcal{P}A, \varnothing, \subseteq, \cap, \cup, - \rangle$$

which is closed under the operations, but not their generalizations. To see this, let $A = \{\{\varnothing\}\}$. Then

$$\mathcal{P}A = \{\varnothing, \{\{\varnothing\}\}\}$$
$$\cup A = \{\varnothing\}$$

so $\cup A \notin \mathcal{P}A$.

2.3 FUNCTIONS AND RELATIONS

2.3.1 ORDERED PAIRS

We now begin introducing a rigorous way of defining functions and relations. Set theory only considers one basic binary relation: membership. Other relations do exist (some argue not), so how can we represent them using sets. One way is to think of extensions for relations. The elements of a relation are clearly ordered ($2 \geq 1$, but $1 \ngeq 2$).

To represent order like this in sets, we define an **ordered pair**, which allows us to get directionality.

Definition 2.1 (Kuratowski's Definition of an Ordered Pair)**.** *We define the* ***ordered pair*** *of* x, y*, as*

$$\langle x, y \rangle := \{\{x\}, \{x, y\}\}$$

This gives us directionality.

Proposition 2.2. $\langle x, y \rangle = \langle u, v \rangle \iff x = u \wedge y = v$

Proof. The $\implies$ direction is trivial, so suppose that $\{\{x\}, \{x, y\}\} = \{\{u\}, \{u, v\}\}$. We have two cases:

1. If $x = y$, then
$$\{\{x\}, \{x, y\}\} = \{\{x\}\} = \{\{u\}, \{u, v\}\}$$
Since $\{\{x\}\}$ is a singleton, so too must $\{\{u\}, \{u, v\}\}$, meaning $u = v$, and
$$\{\{u\}, \{u, v\}\} = \{\{u\}\}$$
It follows that $x = y = u = v$.
2. If $x \neq y$, then $u \neq v$ (otherwise we can redo the first case). Since the sets are equal, $\{x\} \in \{\{u\}, \{u, v\}\}$, so either $\{x\} = \{u\}$, or $\{x\} = \{u, v\}$. The second case is impossible as $\{u, v\}$ is not a singleton. Thus $x = u$, and $\{u, y\} \in \{\{u\}, \{u, v\}\}$. So $\{u, y\} = \{u, v\}$, so $y = v$.

□

Remark. *The definition* $\langle x, y \rangle := \{x, \{x, y\}\}$ *does work but only if we assume the axiom of regualarity.*

2.3.2 CARTESIAN PRODUCTS, BINARY RELATIONS, AND GENERALIZATIONS

We can define a special class whose elements are ordered pairs, with elements of the pairs coming from classes. We first need a lemma:

Lemma 2.3. *If* $x, y \in A$*, then* $\langle x, y \rangle \in \mathcal{P}\mathcal{P}A$.

Proof. If $x, y \in A$, then

$$\begin{aligned} \{x\}, \{x, y\} \subseteq A &\implies \{x\}, \{x, y\} \in \mathcal{P}A \\ &\implies \{\{x\}, \{x, y\}\} \subseteq \mathcal{P}A \\ &\implies \{\{x\}, \{x, y\}\} \in \mathcal{P}\mathcal{P}A \end{aligned}$$

□

Corollary 1. *For any sets A, B, the class*

$$\{u | \exists x, y[x \in A \wedge y \in B \wedge u = \langle x, y \rangle]\}$$

is a set.

Proof. Apply separation to the set $\mathcal{P}\mathcal{P}(A \cup B)$. □

Definition 2.4. *The class created in the above corollary is called the **Cartesian product** of A and B, denoted $A \times B$.*

We are now ready to define what a general binary relation is.

Definition 2.5. *A class A is a **binary relation** if and only if all elements are ordered pairs. That is*

$$\forall x[x \in A \implies \exists u \exists v(x = \langle u, v \rangle)]$$

*For a class A, the **domain** of A, $\operatorname{dom} A$, the **range** of A, $\operatorname{ran} A$, and the **field** of A, $\operatorname{fld} A$, are*

$$\operatorname{dom} A = \{x | \exists y(\langle x, y \rangle \in A)\}$$

$$\operatorname{ran} A = \{x | \exists u(\langle u, x \rangle \in A)\}$$

$$\operatorname{fld} A = \{x | x \in \operatorname{dom} A \vee x \in \operatorname{ran} A\}$$

If A is a set, so too is its domain, range, and field, sinice the domain and range are subsets of $\cup(\cup A)$. Given a relation R, we denote its ordered pairs $\langle x, y \rangle \in R$ as Rxy or xRy.

The concept of an ordered pair and binary relation may also be genearlized to ordered n-tuples and n-ary relations. These are defined inductively:

Definition 2.6. *An ordered $k+1$-tuple $\langle x_1, \ldots, x_k, x_{k+1} \rangle$ is an ordered pair*

$$\langle \langle x_1, \ldots, x_k \rangle, x_{k+1} \rangle$$

where $\langle x_1, \ldots, x_k \rangle$ is an ordered k-tuple. An ordered pair is an ordered 2-tuple. The components of these pairs are called coordinates.

An n-ary relation is a class whose elements are n-tuples, and an n-ary relation on a set A is the a set of n-tuples whose components are in A.

2.3.3 FUNCTIONS

Definition 2.7. *A **function** F is a binary relation such that*

$$\forall x[x \in \operatorname{dom} F \implies \forall y \forall z(Fxy \wedge Fxz \implies y = z)]$$

This just means that each element of the domain maps to a unique object; we call this object for which Fxy for f **the value of** F **at** x, and is denored $F(x)$.

Functions can also have special properties:

Definition 2.8. *F a function from A **into** B if $\operatorname{dom} F = A$ and $\operatorname{ran} F \subseteq B$. We write*

$$F : A \to B$$

*If $\operatorname{ran} F = B$, then F is **surjective**, or **onto** B.*

*We say F is **injective** or **one-to-one** if and only if for each $y \in \operatorname{ran} F$ there is a unique $x \in \operatorname{dom} F$ such that $F(x) = y$. In general, if a set R satisfies this condition is **single rooted**.*

*If F is both injective and surjective then it is a **bijection**.*

Given sets A, F, G, then the separation axiom allows us to define other sets related to a function.

Definition 2.9. *The **inverse** F^{-1} of F is the set*

$$F^{-1} = \{\langle v, u\rangle \mid \langle u, v\rangle \in F\}$$

*The **composition** of F, G is the set*

$$F \circ G = \{\langle u, v\rangle \mid \exists x(Gux \wedge Fxv)\}$$

*The **restriction** of F to A is the set*

$$F \restriction A = \{\langle u, v\rangle \mid Fuv \wedge u \in A\}$$

*The **image** of A under F is the set*

$$F[A] = \operatorname{ran}(F \restriction A)$$

*The **inverse image** of A under F is the set*

$$F^{-1}[A] = \{x \in \operatorname{dom} F | F(x) \in A\}$$

2.4 THE AXIOM OF CHOICE (AC)

2.4.1 FIRST FORM

One can prove many results pertaining to the concepts described above. One of which is the following:

Theorem 2.10. *Suppose that $F : A \to B$ with $A \neq \varnothing$.*

(a) *There is a function $G : B \to A$, called the left inverse, such that $G \circ F = I_A$ if and only if F is an injection.*

(b) *There is a function $H : B \to A$,called the right inverse, such that $F \circ H = I_B$ if and only if F is surjective.*

Note that $I_C = \{\langle z, z\rangle \,|z \in C\}$ for any C.

The problem here is the $\Longleftarrow$ direction in (b). Suppose that F is surjective. Then for all $y \in B$, we have that

$$F^{-1}[\{y\}] \neq \varnothing$$

We know F^{-1} is a relation whose domain is B and range is A, but we can't say its a function, because $F^{-1}[\{y\}]$ is not necessarily a singleton. What we need is some subset of F^{-1}, H that is single rooted. The Axiom of Choice guarantees this:

The Axiom of Choice (First Form): For any relation R there is a single rooted subset S of R such that $\operatorname{dom} S = \operatorname{dom} R$.

From here we can prove our theorem.

Proof. Let H be the single rooted subset of F^{-1} guaranteed by AC. Clearly $H : B \to A$ and $\langle y, x\rangle \in H$ if and only if $\langle x, y\rangle \in F$. Thus, $\langle u, v\rangle \in F \circ H$ if and only if there is an $x \in A$ such that $\langle u, x\rangle \in H$ and $\langle x, v\rangle \in F$. But as $H \subseteq F^{-1}$, we have that $\langle x, u\rangle \in H$ and so $u = v$, meaning $F \circ H = I_B$. □

This just means you are able to "choose" an element of each preimage in constructing your inverse (this is why its called the Axiom of *Choice*).

2.4.2 SECOND FORM

An alternate form of AC may be constructed using genearlized Cartesian products. Let R be afunction and $I \subseteq \operatorname{dom} R$. We consider

$$X_{i\in I}R(i) = \{f|f \text{ is a function and } \operatorname{dom} f = I \wedge \forall i \in I(f(i) \in R(i))\}$$

You essentially start with a bunch of sets $R(i)$, and this new abstract we've defined is the set of functions from I to the union of $R(i)$'s, where we choose a point in $R(i)$ that i is mapped to by f. Given I just used the word "choose," you wouldn't be wrong to assume AC is relevant here.

The Axiom of Choice (Second Form): For any I and function R with $\operatorname{dom} R = I$, if $R(i) \neq \varnothing$ for all $i \in I$, then $X_{i\in I}R(i) \neq \varnothing$.

It is a good exercise to show that these two forms of AC are equivalent.

3

Lecture 3: Constructing the Naturals

In this lecture, we will see for the first time how Set Theory acts as a framework for mathematics; in particularly, we will see how it is a framework for number theory.

In *Principa Mathematica*, the authors famously prove $1 + 1 = 2$, hundreds of pages into the book, using very formal mathematics. It is a somewhat humorous proof, given the intuitive nature of the result. But, someone might look at that proof and ask "What is 1? What is 2? What is $+$?" It was Frege who first asked these questions in this context. Now, we are going to follow in their footsteps and answer these questions.

Thus far, we have developed some basic tools, mainly axioms, and some machinery related to relations and functions. At this point, we can now reconstruct some key mathematical structures within set theory. These days, such a structure consists of a set (the domain) along with relations, functions, and distinguished elements. The most important one is the structure of the **natural numbers**. Frege and Dedekind's work in set theory was largely justfied by a desire to find a philosophical conception of natural numbers. The relationship between number theory (arithmetic) and set theory remains an open question, and it is useless to say one is more fundamental than the other. With that said, a famous result of Gödel says that while ZFC is sufficient in develop a good theory of number systems like the integers, reals, etc., some simple arithmetic questions cannot be settled in ZFC.

3.1 THE EXISTENCE OF THE NATURALS

3.1.1 SIMPLE INFINITE SYSTEMS AND THE AXIOM OF INFINITY

What are natural numbers? Where does our knowledge of them come from? Dedekind had an abstractionist view of numbers. He defines a **simple infinite system** M as a set/system satisfying

(i) M is closed under an injective function f

(ii) There is an $e \in M$ such that $e \notin \operatorname{ran} f$

(iii) If X satisfies (i) and (ii) then $M \subseteq X$.

He says that f **orders** M and that e is a **base element** for M with respect to f. He then defines the natural numbers as such a simple infinite system:

If, in considering a simply infinite system N ordered by the mapping ϕ, we completely disregard the particular nature of the elements, retaining only their distinguishability and considering

only those relationships in which they are placed to one another by the ordering map ϕ, then these elements are called natural numbers or ordinal numbers or simply numbers, and the base element e is called the base element of the number series N. In consideration of this freeing of the elements from every other content (abstraction) one can with justice call the numbers a free creation of the human intellect(menschlichen Geistes)

This definition still does not answer the question of if a simple infinite system exists. While Dedekind tried to prove their existence, it is widely regarded to be a failure. Other philosophers, such as Parsons, have tried to use other methods to prove its existence, but have failed to do so. Thus, instead of trying to prove that one exists, let's just assume that it does using an axiom.

Definition 3.1. *For any set a, the **successor** a^+ is the set $a \cup \{a\}$.*

*We say A is an **inductive set** if and only if $\varnothing \in A$, and A is closed under the set successor operation.*

The Axiom of Infinity: There is an inductive set A.

Definition 3.2. *The set of natural numbers ω is the intersection of all inductive sets.*

We know that ω exists thanks to the existence of A. We can attain another inductive set $A \cup \{A^+\}$. This gets us many more inductive sets, and by applying the separation schema to get a subset of an inductive set containing only the elements that are in all other inductive sets that exists, we get ω.

It should be evident that, as $\varnothing$ is in all inductive sets, it is in ω. Furthermore, If $a+ \notin \omega$ for some set a, then a would be in all inductive sets, but a^+ would not, contradicting the fact that all inductive sets are closed under set successor. Thus,

Theorem 3.3. *ω is inducitve, and coincides with any inductive subset of it.*

This means that if we have a subset of ω that is itself inductive, we can find a correspondence between it and ω. From this, we get a fundamental principle

Corollary 2 (Principle of Induction)**.** *If $A \subseteq \omega$, and A is inductive, meaning $\varnothing \in A$ and is closed under set successor, then $A = \omega$.*

This is the main principle we will use in our development of arithmetic.

3.1.2 THE DEDEKIND-PEANO POSTULATES

Now that we know the natural numbers exist, it's time to use it to formalize arithmetic in a set theoretic context. We start by considering systems that generalize the properties of the naturals, mainly that there is a 0 element, the successor function is injective, and the principle of induction.

Definition 3.4 (Dedekind-Peano Postulates)**.** *A **Peano system** is a triple $\langle N, S, e\rangle$ with N a set, $S : N \to N$, and $e \in N$ such that*

(i) $e \notin \operatorname{ran} S$

(ii) S *is injective*

(iii) If $A \subseteq N$, $e \in A$, and $S[A] \subseteq A$, then $A = N$

There is another special property of the natural numbers that we need to define as well: any element of the naturals is also a subset of the naturals:

Definition 3.5. *A is **transitive** if and only if A satisfies*

$$\forall x \forall y[(x \in y \wedge y \in A) \implies x \in A]$$

Equivalently, A satifies the property that

$$\cup A \subseteq A \quad or \quad x \in A \implies x \subseteq A \quad or \quad A \subseteq \mathcal{P}A$$

Lemma 3.6. *(i) If a is transitive, then $\cup(a^+) = a$*

(ii) Every natural number is a transitive set

(iii) ω is transitive

Proof. (i) For a transitive set a,

$$\begin{aligned} \cup a^+ &= \cup(a \cup \{a\}) \\ &= (\cup A) \cup (\cup\{a\}) \\ &= (\cup a) \cup a \\ &= a \qquad \text{(by transitivity)} \end{aligned}$$

(ii) We prove it by induction. Let

$$T = \{n \in \omega | n \text{ is a transitive set}\}$$

We claim that $T = \omega$. $0 \in T$ trivially. Moreover, if $k \in T$, then by (i)

$$\cup(k^+) = k \subseteq k^+$$

so $k^+ \in T$, and we conclude that T is inductive. By induction, $T = \omega$.

(iii) Again we prove it by induction. Let

$$T = \{n \in \omega | n \subseteq \omega\}$$

We claim $T = \omega$. Clearly $0 \in T$. Moreover, if $k \in T$, then $k \subseteq \omega$ and $\{k\} \subseteq \omega$. Thus $k \cup \{k\} \subseteq \omega$ so $k^+ \in T$. By induction we conclude that $T = \omega$ as T is inductive.

□

With this lemma we can create a Peano system using the naturals:

Theorem 3.7. *$\langle \omega, \sigma, 0 \rangle$ is a Peano system, where σ is the restriction of set theoretic successor to ω.*

Proof. Clearly $0 \notin \operatorname{ran}\sigma$. Moreover if $A \subseteq \omega, 0 \in A$, and $\sigma[A] \subseteq A$, then $A = \omega$. Finally, if $m^+ = \sigma(m) = \sigma(n) = n^+$, then by the above Lemma,

$$\cup m^+ = \cup n^+ = m = n$$

so σ is injective. □

We have thus formalized the natural numbers in terms of set theory. This is progress, but we're not done yet. While we've gotten the elements down, we still need operations and relations.

First off, operations. Where does + come from in a set theoretic context? This is answered by the Recursion Theorem. Suppose I gave you a function $h : \omega \to A$ and a function $F : A \to A$ such that $h(0)$ is given and $h(n^+) = F(h(n))$. Then we can find out what h is by computing each value successively:

$$h(0), h(1) = F(h(0)), h(2) = F(h(1)) = F(F(h(0))), \ldots$$

In other words, $h(m) = F^m(a)$, where F^m just means repeating F m times.

We can show that for any set A, $a \in A$, and map $F : A \to A$, that such a function h exists.

Theorem 3.8 (The Recursion Theorem)**.** *If A is a set $a \in A$, and $F : A \to A$, then there is a unique function $h : \omega \to A$ such that*

(i) $h(0) = a$,

(ii) $h(n^+) = F(h(n))$ *for all* $n \in \omega$.

To prove it we need a new definiton. A function h_k is called **acceptable** if and only if its domain is a subset of ω, its range a subset of A, and

(i) $0 \in \operatorname{dom} v \implies v(0) = a$

(ii) $n^+ \in \operatorname{dom} v \implies n \in \operatorname{dom} v \wedge v(n^+) = F(v(n))$

Proof. Let h be the union of all acceptable functions. In other words, $h(n) = y$ if and only if $v(n) = y$ for an acceptable function v.

First we show h is a function. It suffices to prove that two acceptable functions agree on the intersection of their domains. Let

$$S = \{n \in \omega | h(n) = y \text{ for at most one } y\}$$

We claim S is inductive and thus ω. If

$$y_1 = h(0) = y_2$$

for some y_1, y_2 then there exists v_1, v_2 such that $v_1(0) = y_1, v_2(0) = y_2$, so $y_1 = a = y_2$. So $0 \in S$.

Now suppose that $k \in S$. If

$$y_1 = h(k^+) = y_2$$

for some y_1, y_2 then there exists v_1, v_2 such that $v_1(k^+) = y_1, v_2(k^+) = y_2$. Thus,

$$y_1 = F(v_1(k)) \quad y_2 = F(v_2(k))$$

and because $k \in S$, we know that $v_1(k) = h(y) = v_2(k)$, so $F(v_1(k)) = F(v_2(k))$. Thus, $k^+ \in S$, and so by induction $S = \omega$.

h is a function, now we need to show it is acceptable. It is clear from the definition that the domain and ranges are ω and A respectively. We just need to check the other conditions.

For (i), if $0 \in \operatorname{dom} h$, then there is an acceptable v such that $v(0) = h(0)$. As $v(0) = a$, $h(0) = a$, as desired.

For (ii), if $n^+ \in \operatorname{dom} h$, then there is an acceptable v such that $v(n^+) = h(n^+)$. As v is acceptable, we have that $n \in \operatorname{dom} v$ and since $v(n) = h(n)$,

$$h(n^+) = v(n^+) = F(v(n)) = F(h(n))$$

so h is indeed acceptable.

We now show that $\operatorname{dom} h = \omega$. We show that $\operatorname{dom} h$ is inductive. As $\{\langle 0, a\rangle\}$ is acceptable, $0 \in \operatorname{dom} h$. Now suppose $k \in \operatorname{dom} h$, and suppose for the sake of contradiction that $k^+ \notin \operatorname{dom} h$. Then consider

$$v = h \cup \{\langle k^+, F(h(k))\rangle\}$$

Then v is a function whose domain and ranges satisfy the requirements for an acceptable function. Moreover, $v(0) = h(0) = a$. In addition, if $n^+ \in \operatorname{dom} v$, with $n^+ \neq k^+$, then $n^+ \in \operatorname{dom} h$ and $v(n^+) = h(n^+) = F(h(n)) = F(v(n))$. Similarly if $n^+ = k^+$, by injectivity of the successor function, $n = k$, and as $k \in \operatorname{dom} h$,

$$v(k^+) = F(h(k)) = F(v(k))$$

so v is an acceptable function. But this means that $v \subseteq h$, and so $k^+ \in \operatorname{dom} h$, a contradiction. Thus, $\operatorname{dom} h$ is inducitve and so by induction it is equal to ω.

Finally, we claim that h is unique. Let h_1, h_2 satisfy the theorem. Let

$$S = \{n \in \omega | h_1(n) = h_2(n)\}$$

we claim S is inductive. Clearly $0 \in S$ as $h_1(0) = a = h_2(0)$ by construction. Now, assume $k \in S$ and consider k^+. We have that

$$h_1(k^+) = F(h_1(k)) = F(h_2(k)) = h_2(k^+)$$

so $k^+ \in S$. So S is inductive and by induction $S = \omega$, so $h_1 = h_2$. □

There are many other Peano systems; the system where N is the powers of 2, S is multiplication by 2, and $e = 1$ is a Peano system. However, one can readily see that this system is equivalent to $\langle \omega, \sigma, 0\rangle$. In fact, any Peano system is essentially the same as $\langle \omega, \sigma, 0\rangle$!

Theorem 3.9 (The Isomorphism Theorem)**.** *$\langle \omega, \sigma, 0\rangle$ is isomorphic to any Peano system $\langle N, S, e\rangle$. There is a bijection $h : \omega \to N$ such that*

(i) $h(0) = e$

(ii) $h(\sigma(n)) = S(h(n))$

and we denote this by $\langle \omega, \sigma, 0\rangle \cong \langle N, S, e\rangle$.

Proof. Let h be the function guaranteed from the Recursion Theorem. We claim that h is injective and it's range is N. This would show thta it is a bijection satifying the requirements.

First we show that $\operatorname{ran} h = N$. We use induction on $\langle N, S, e\rangle$. Clearly $e \in \operatorname{ran} h$ as $h(0) = e$. Furthermore, for any $x \in \operatorname{ran} h$ where $x = h(n)$, we have that $h(n^+) = S(x)$, so $S(x) \in \operatorname{ran} h$. Thus by induction $\operatorname{ran} h = N$.

We now show h is injective. Let

$$T = \{n \in \omega | \text{for all } m \in \omega \text{ with } m \neq n,\ h(m) \neq h(n)\}$$

We claim T is inductive. Any $m \neq 0$ must be the successor of some value p, and $h(p^+) = S(h(p)) \neq e$ as e isn't in the range, so $0 \in T$.

Now suppose $k \in T$ and consider k^+. Suppose $h(k^+) = h(m)$. Then from before we know that $m \neq 0$, so $m = p^+$ for some p, and we get

$$S(h(k)) = h(k^+) = h(p^+) = S(h(p))$$

S is injective, so we have that $h(k) = h(p)$. As $k \in T, k = p$, so $k^+ = p^+ = m$. Thus, $k^+ \in T$, so T is inductive and we conclude that $T = \omega$, hence h is injective. □

Because of these theorems, the theory whose axioms are the Dedekind-Peano postulates, called **second-order (Peano) arithmetic**, is a categorical theory.

3.2 THE ELEMENTARY THEORY OF ARITHMETIC

3.2.1 ADDITION AND MULTIPLICATION

It is now relatively straightforward to define and prove some basic properties of addition and multiplication:

Definition 3.10. *Given any $m \in \omega$, The Recursion Theorem gives us a unique function*

$$A_m : \omega \to \omega$$

such that $A_m(0) = m$, $A_m(n^+) = A_m(n)^+$ for every $n \in \omega$.

*We thus define **addition** $+$ to be the function on $\omega \times \omega$ such that for $m, n \in \omega$,*

$$m + n = A_m(n)$$

Definition 3.11. *By the Recursion Theorem, there is a unique function*

$$M_m : \omega \to \omega$$

*such that $M_m(0) = 0$ and $M_m(n^+) = M_m(n) + m$ for every $m, n \in \omega$. We then define **multiplication** $\cdot$ as the function for which $m \cdot n = M_m(n)$.*

Exponentiation can be done in a similar fashion. From here, we are then able to state and prove many elementary properties of addition and multiplication.

Theorem 3.12. *The following equations hold for all $m, n, p \in \omega$:*

(i) $m + 0 = m$

(ii) $m + n^+ = (m + n)^+$

(iii) $m \cdot 0 = 0$

(iv) $m \cdot n^+ = m \cdot n + m$

(v) $(m + n) + p = m + (n + p)$

(vi) $m + n = n + m$

(vii) $m \cdot (n + p) = m \cdot n + m \cdot p$

(viii) $(m \cdot n) \cdot p = m \cdot (n \cdot p)$

(ix) $m \cdot n = n \cdot m$

3.2.2 ORDERING THE NATURALS

While we have defined operations on the naturals, we have yet to define relations on them, mainly the relations of $<$ and $\leq$. In the context of our development of arithmetic, we define $m < n$ through membership, meaning that

$$m < n \iff m \in n$$

Similarly, $m \leq n$ if and only if $m < n \vee m = n$. Clearly, $m < n \iff m^+ \leq n$. This leads to an important theorem about ordering the naturals:

Theorem 3.13 (Trichotomy of the Order Relation on ω)**.** *For every $m, n \in \omega$, exactly one of the below holds:*

$$m \in n \quad n \in m \quad m = n$$

In other words, either $m < n$, $n < m$, or $m = n$. Many corollaries follow from the Trichotomy:

Theorem 3.14. *The following hold for all $m, n \in \omega$*

(i) $m \in n \iff m \subset n$

(ii) $m < n \iff m + p < n + p$ *for all* $p \in \omega$

(iii) $m < n \iff m \cdot p < n \cdot p$ *for all* $p \in \omega$ *with* $p \neq 0$

(iv) $m + p = n + p \implies m = n$ *for all* $p \in \omega$

(v) $m \cdot p = n \cdot p \implies m = n$ *for all* $p \in \omega$ *with* $p \neq 0$

We can also use the Trichotomy to prove two other powerful principles that are other forms of induction:

Theorem 3.15 (The Well-Ordering Principle)**.** *If $A \subseteq \omega$ and $A \neq \varnothing$, then there is an $n \in A$ such that $n \leq m$ for all $m \in A$ (Every subset of the naturals has a smallest element).*

Proof. Let $A \subseteq \omega$ and suppose A has no least smallest element. We claim $A = \varnothing$.

Let $B = \{m \in \omega : \forall n(n \in m \implies m \notin A)\}$. We claim B is inductive. $0 \in B$ vacuously holds since there are no members of 0. Now suppose $m \in B$ and consider m^+. Suppose by way of contradiction that $n \in m^+$ and $n \in A$. Then by Trichotomy $n \in m$ or $n = m$. $n \notin m$ as $m \in B$, so $n = m$. But as $m \in B$ and $n \in A$, no element smaller than n is in A. This means n is A's smallest element, a contradiction.

B is thus inductive. Now, if A was nonempty, there is some $n \in A$. This means $n^+ \notin B$, which contradicts that $B = \omega$. □

Theorem 3.16 (The Principle of Strong Induction)**.** *If $A \subseteq \omega$ and*

$$\forall m(m < n \implies m \in A) \implies n \in A$$

for all $n \in \omega$, then $A = \omega$ (If A contains every number smaller than n for all n, then A is the naturals).

Remark. *Throughout this lecture, we have built up the theory of second-order arithmetic, including natural numbers, the operations of addition and multiplication, and the orderings $<$ and $\leq$, in the language of ZF without the replacement axiom. We have thus shown that ZF without the replacement axiom, can* ***interpret*** *second-order arithmetic. In particular, we have shown that we can translate every primitive symbol in second-order arithmetic, its elements, functions, and relations, into the language of ZF. Moreover, we can show that for every axiom of the second-order arithmetic, we can translate it into ZF and in fact prove it. We thus write*

$$ZF \succeq PA^2$$

More interestingly, as stated in Lecture 2, we can define sets using numbers and the theories of sets may be interpreted in arithmetical theories. If we let ZF^- be ZF where the axiom of infinity is replaced with its negation, then

$$ZF^- \equiv PA$$

Where $\equiv$ is the same as saying $\succeq$ and $\preceq$ at the same time. As another example of these equivalencies, if we let Q be Peano arithmetic without induction, and UST the set theory with only the empty set and pairing axioms, then

$$Q \equiv UST$$

4

Lecture 4: Number Systems

After constructing the natural numbers, we continue constructing new number systems by building the integers, rationals, and the reals.

Remark. *We're not going to construct the complex numbers, as* $\mathbb{C} = \mathbb{R} \times \mathbb{R}$.

We also will not go into depth on the hyperreals, which is $\mathbb{R}$ *alongside infinitesimals and infinitely large numbers.*

4.1 THE INTEGERS

To construct the integers, we need to define what a negative number is. All negative numbers are the difference of two positive numbers. For instance, $2 - 4 = -2$, and $5 - 8 = -3$. A naïve definition of a negative number can thus be an ordered pair.

$$-2 := \langle 2, 4 \rangle$$

But this leads to a problem. Under this definition, we'd get that

$$-2 = \langle 0, 2 \rangle = \langle 1, 3 \rangle = \langle 3, 5 \rangle = \cdots$$

in fact, there are an infinite number of such ordered pairs that would equal -2. What we need to do is to collapse this infinite set into a single object that we can say is -2. Equivalence classes are perfect for this!

Definition 4.1. *We define the relation* $\sim$ *on* $\omega \times \omega$ *as follows:*

$$\langle m, n \rangle \sim \langle p, q \rangle \iff m + q = p + n$$

This definition is natural; we want $m - n = p - q$, which is equivalent to the above by placing q on the left and n on the right.

Theorem 4.2. $\sim$*, as defined above, is an equivalence relation.*

Proof. Reflexivity is obvious since $m + n = m + n$. For symmetry, suppose $\langle m, n \rangle \sim \langle p, q \rangle$. Then by commutativity of addition,

$$m + q = p + n \iff p + n = m + q$$

so $\langle p, q\rangle = \langle m, n\rangle$.

For transitivity, suppose $\langle m, n\rangle \sim \langle p, q\rangle$ and $\langle p, q\rangle \sim \langle r, s\rangle$. Then we have that

$$m + q + p + s = p + n + r + q$$

Cancelling $q + p$ on both sides gives us

$$m + s = r + n$$

and so $\langle m, n\rangle \sim \langle r, s\rangle$. □

Definition 4.3. *The **integers** $\mathbb{Z}$ is the set $(\omega \times \omega)/\sim$, the set of ordered pairs modulo $\sim$.*

For example

$$2_Z = [\langle 2, 0\rangle] = \{\langle 2, 0\rangle, \langle 3, 1\rangle, \ldots\}$$

$$-3_Z = [\langle 0, 3\rangle] = \{\langle 0, 3\rangle, \langle 1, 4\rangle, \ldots\}$$

We should quickly check that $\mathbb{Z}$ is actually a set. Indeed, if $x \in [\langle m, n\rangle]$, then

$$\phi(x) = \exists k \exists l[k \in \omega, l \in \omega, x = \langle k, l\rangle, \langle k, l\rangle \sim \langle m, n\rangle]$$

holds. From here, we see that

$$[\langle m, n\rangle] = \{x \in \omega \times \omega : \phi(x)\}$$

which is a set by the separation schema. Finally $\mathbb{Z}$ has these sets as elements, so $\mathbb{Z} \in \mathcal{P}(\omega \times \omega)$, and thus is a set.

We must now give $\mathbb{Z}$ an addition operation. Intuitively we should be able to add differences,

$$(m - n) + (p - q) = (m + p) - (n + q)$$

so we will define our addition operation $+_Z$ as

$$[\langle m, n\rangle] +_Z [\langle p, q\rangle] := [\langle m + p, n + q\rangle]$$

but we're not done. Recall that we are working with equivalence classes. To show that $+_Z$ is indeed the correct addition operator, we need to show that the value does not change if we use a different representative of a class:

Lemma 4.4. *If $\langle m, n\rangle \sim \langle m', n'\rangle$ and $\langle p, q\rangle \sim \langle p', q'\rangle$, then*

$$\langle m + p, n + q\rangle \sim \langle m' + p', n' + q'\rangle$$

Proof. Follows from commutativity and associativity of $+$. □

The properties of $+$ also extend to $+_Z$. In particular,

Theorem 4.5. *$+_Z$ is commutative and associative.*

Proof. Straightforward □

$\mathbb{Z}$ with $+_Z$ has a group structure. In particular, it has an identity element 0_Z and inverses:

Theorem 4.6. *$0_Z = \langle 0, 0\rangle$ is the identity element of $+_Z$: for any $a \in \mathbb{Z}$, $a +_Z 0_Z = a$. Moreover, there is an integer b such that $a +_Z b = 0_Z$, and this b is unique*

Proof. The first claim is obvious. For the second claim, we let $b = [\langle n, m\rangle]$. Then

$$a +_Z b = [\langle m + n, n + m\rangle] = [\langle 0, 0\rangle] = 0_Z$$

as desired. For uniqueness, let b, b' be inverses of a. Then observe that

$$b = b +_Z (a +_Z b') = (b +_Z a) +_Z b' = b'$$

as desired. □

As inverse are unique, we can denote the inverse of a to be $-a$. These give us a subtraction operation, defined as

$$b - a := b +_Z (-a)$$

We are also able to give $\mathbb{Z}$ a multiplication operation, which we define in a very similar way to addition. We have

$$(m - n) \cdot (p - q) = (mp + nq) - (mq + np)$$

Thus, we define $\cdot_Z$ as follows:

$$[\langle m, n\rangle] \cdot_Z [\langle p, q\rangle] := [\langle mp + nq, mp + np\rangle]$$

Again, we'd need to verify that this is well-defined and does not change value if we use other representatives, a proof we omit. We can also prove some results analagous to those for addition:

Theorem 4.7. *$\cdot_Z$ is commutative, associative, and distributive over $+_Z$.*

Theorem 4.8. *(i) The integer $1_Z = [\langle 1, 0\rangle]$ is the multiplicative identity element: for all $a \in \mathbb{Z}$, $a \cdot_Z 1_Z = a$.*

(ii) $0_Z \neq 1_Z$

(iii) If $a \cdot_Z b = 0_Z$, then $a = 0_Z$ or $b = 0_Z$.

Example 1. *One can show that*

$$[\langle 0, 1\rangle] \cdot_Z [\langle m, n\rangle] = [\langle n, m\rangle]$$

telling us that $-1_Z \cdot_Z a = -a$.

Remark. *All of theses results combine to tell us that $\mathbb{Z}$ is a **integral domain**, a concept that shows up in abstract algebra and ring theory.*

We also need to develop some sort of ordering of the integers. Again, this follows from the ordering on the naturals. Recall that

$$m - n < p - q \iff m + q < p + n$$

Thus, we will define our ordering $<_Z$ as follows:

$$[\langle m, n\rangle] <_Z [\langle p, q\rangle] \iff m + q \in p + n$$

Like addition and multiplication, $<_Z$ is well-defined under equivalence classes (we again omit the proof).

Theorem 4.9. *$<_Z$ is a linear ordering on $\mathbb{Z}$, meaning it is transitive and satisfies the trichotomy on $\mathbb{Z}$.*

Definition 4.10. *An integer b is **positive** if $0_Z <_Z b$.*

Equivalently, one can say that $0 <_Z -b$ (this is easy to prove). A consequence of trichotomy is thus that for any integer b, it is either positive, negative, or zero.

Theorem 4.11. *For any $a, b, c \in \mathbb{Z}$,*

(i) $a <_Z b \iff a +_Z c <_Z b +_Z c$

(ii) If $0_Z <_Z c$, *then* $a <_Z b \iff a \cdot_Z c <_Z b \cdot_Z c$.

4.2 THE RATIONALS

From now on we drop the subscript on elements and operations on the integers. We now need to construct fractions, the set of which is the rational numbers. The obvious idea is to define a rational number, say 2/3, as a fraction. In other words,, it is an ordered pair $\langle a, b\rangle$ where $b \neq 0$. We call b the **denominator** and a the **numerator**. Like the integers, this definition can lead to different ordered pairs equalling the same fraction. We can avoid this by again defining an equivalence relation: Let $\mathbb{Z}'$ be $\mathbb{Z}$ without 0.

Definition 4.12. *We define $\sim$ to be the binary relation on $\mathbb{Z} \times \mathbb{Z}'$ satisfying*

$$\langle a, b\rangle \sim \langle c, d\rangle \iff a \cdot d = c \cdot d$$

Theorem 4.13. *$\sim$ is an equivalence relation on $\mathbb{Z} \times \mathbb{Z}'$.*

Proof. We just prove transitivity, as the remaining items are easy to prove by commutativity and associativity of multiplication on the naturals. Suppose $\langle a, b\rangle \sim \langle c, d\rangle$ and $\langle c, d\rangle \sim \langle e, f\rangle$. Then

$$ad = cb \quad \text{and} \quad cf = ed$$

Take the first equation and multiply by f, then take the second and multiply it by b. We get

$$adf = bcf \quad \text{and} \quad cfb = edb$$

Thus, $adf = edb$, and cancelling the nonzero d, we get $af = eb$. Thus, $\langle a, b\rangle \sim \langle e, f\rangle$. □

Definition 4.14. *The **rationals** $\mathbb{Q}$ is the set $(\omega \times \omega)/\sim$.*

We define the constants $0_Q = \langle 0, 1\rangle$ and $1_Q = \langle 1, 1\rangle$. Most of the proofs for rationals follow from proofs of similar results for the integers, so we will exclude them here and invite the reader to give them a try.

Our intuitive notions of addition and multiplication on rationals guides our formal definitions:

$$a/b + c/d = (ad + cb)/bd \quad a/b \cdot c/d = ac/bd$$

Thus, we define

$$[\langle a, b\rangle] +_Q [\langle c, d\rangle] := [\langle ad + cb, bd\rangle]$$
$$[\langle a, b\rangle] \cdot_Q [\langle c, d\rangle] := [\langle ac, bd\rangle]$$

both of which are well-defined under equivalence classes. These operations also work in the ways we want:

Theorem 4.15. *(i) $+_Q, \cdot_Q$ are commutative and associative. $\cdot_Q$ is distributive over $+_Q$.*

(ii) 0_Q is an additive identity for $+_Q$, and 1_Q is a multiplicative identity for $\cdot_+$.

(iii) For all $r \in \mathbb{Q}$, there is an $s \in Q$ such that $r +_Q s = 0_Q$. For all $r \neq 0_Q$, there is a $q \neq 0_Q$ such that $r \cdot_Q q = 1_Q$.

This additive and multiplicative inverses are unique, and we denote them as $-r$ and r^{-1} respectively. We have subtraction as in the integers, as well as division, which we define by

$$s \div r := s \cdot_Q r^{-1}$$

We also need to define an ordering on the rationals. Informally we know that

$$\frac{a}{b} < \frac{c}{d} \iff ad < cb$$

whenever b, d are positive, but this is not always guaranteed. However, we know that

$$[\langle a, b\rangle] = [\langle -a, -b\rangle]$$

so by the trichotomy on the integers, every rational number can be expressed with a positive denominator. We can then define our ordering

$$[\langle a, b\rangle] <_Q [\langle c, d\rangle] \iff ad < cb$$

Lemma 4.16. *Assume $\langle a, b\rangle \sim \langle a', b'\rangle$ and $\langle c, d\rangle \sim \langle c', d'\rangle$, and also assume that b, b', d, d' are positive. Then*

$$ad < cb \iff a'd' < c'b'$$

Theorem 4.17. *$<_Q$ is a linear ordering on $\mathbb{Q}$, as it is transitive and satisfies the trichotomy.*

We call r **positive** if $0_Q <_Q r$. By trichotomy, for any rational r, either r is positive, $-r$ is positive, or r is zero. From this, we can define the **absolute value** of r, $|r|$, as

$$|r| = \begin{cases} -r & -r \text{ is positive} \\ r & \text{otherwise} \end{cases}$$

Order is preserved by the operations on the rationals.

Theorem 4.18. *For rationals r, s, t,*

(i) $r <_Q s \iff r +_Q t <_Q s +_Q t$.

(ii) If t is positive, then $r <_Q s \iff r \cdot_Q t <_Q s \cdot_Q t$.

Finally, we can prove the basic cancellation laws, which will make certain arithmetic proofs much simpler in future.

Theorem 4.19. *For any rationals s, t, r,*

(i) If $r +_Q t = s +_Q t$, *then* $r = s$

(ii) If $r \cdot_Q t = s \cdot_Q t$ *and* $t \neq 0_Q$, *then* $r = s$.

4.3 THE REALS

Something that surprised early mathematicians, particularly those in Ancient Greece, was that the rationals were not sufficient in expressing all values. Take for instance the diagonal of a square with side length 1. The Pythagoreans showed that this value, $\sqrt{2}$, cannot be a rational number, and thus forced them to go beyond the rationals.

Our previous derivations of simpler number systems like the integers and rationals were based on equivalence classes of ordered pairs, building the new number system from established ones. We cannot achieve a similar construction of the reals using the rationals, because as we will later see, the reals are much larger than the rationals. There are, however, many known ways of constructing the reals. We will use the method most well known as *Dedekind cuts*. While it is not the method that provides the best proofs for the arithmetical properties of the reals, it is the one that provides the simplest definition of them.

Definition 4.20. *A **Dedekind cut** is a subset x of $\mathbb{Q}$ such that*

(i) $\varnothing \neq x \neq \mathbb{Q}$

(ii) If $q \in x$ *and* $r < q$, *then* $r \in x$; *x is closed downwards*

(iii) x has no largest element.

Definition 4.21. *We define the **real numbers** $\mathbb{R}$ to be the set of all Dedekind cuts.*

Note there are no equivalence classes; the reals are just the set of Dedekind cuts. We get operations $+_R$ and $\cdot_R$ which are applied based on their Dedekind cuts (we omit the definition and proofs of their properties). We also get an ordering on $\mathbb{R}$ given by

$$x <_R y \iff x \subset y$$

Theorem 4.22. *$<_R$ is a linear ordering on $\mathbb{R}$.*

This allows us to prove a very important quality of $\mathbb{R}$, one that in fact makes it unique amongst all similar structures.

Definition 4.23. *Let A be a set of reals. A real number x is an **upper bound** of A iff for all $y \in A$,*

$$y \leq_R x$$

*The set A is **bounded (above)** iff there is an uppper bound of A. A **least upper bound** of A is an upper bound that is less than any other upper bound.*

Example 2. *Consider the set $\{r \in \mathbb{Q} | r \cdot r < 2\}$ in $\mathbb{R}$. This set is bounded but has no least upper bound in $\mathbb{Q}$ (this follows from $\sqrt{2}$ being irrational, but we will not prove it).*

The example above can be rectified by considering the set in $\mathbb{R}$. Now it does have a least upper bound, $\sqrt{2}$. This is a general property of subsets of the reals:

Theorem 4.24 (The Least Upper Bound Property)**.** *Any bounded, nonempty subset of $\mathbb{R}$ has a least upper bound in $\mathbb{R}$.*

Proof. Let A be the set in question. We claim that the least upper bound is $\bigcup A$.

By definition of $\bigcup A$, for all $x \in A$, $x \subseteq \bigcup A$. Let z be an upper bound for A, so $x \subseteq z$ for all $x \in A$. Then

$$\bigcup A \subseteq z$$

Thus, $\bigcup A$ is the least upper bound of A with respect to ordering by inclusion. We now show that $\bigcup A$ is a real number. $\bigcup A \neq \varnothing$ since A is nonempty. Moreover, as $\bigcup A \subseteq z$ for an upper bound z of A, $\bigcup A \neq \mathbb{Q}$. Finally, $\bigcup A$ is closed downwards and has no largest element, since if either claim failed, it would fail on an element of A, thus contradicting that A is bounded. $\square$

5

The Size of Sets

We've constructed several different sets. We now study the size of these sets. The size of sets will form the basis for our notion of cardinal numbers and cardinal arithmetic.

5.1 EQUINUMEROSITY

Given two sets A, B, we want to determine when A and B have the same size, or when A has more elements than B. This question is easily answered when these sets are finite, as well as when one is finite and the other is infinite. But what if A and B are both infinite? Now we need to define what it means for two infinite sets have the same size. Maybe we can use the definition that A is smaller than B if $A \subsetneq B$, but this is not satisfactory. Let $B = \mathbb{N}$ and $A = 2\mathbb{N}$, the set of even natural numbers. Both sets have the same size, even though one is contained within the other.

To give a good definition, we consider an analogy: suppose you are in grade school, and are given a box of circles and a box of triangles, and tasked with determining whether or not these boxes have the same size. Because you cannot count higher than 3, and there are clearly more than three of each, you seem stuck. But not all hope is lost. You take out one circle and pair it with one triangle. You repeat this process and are able to pair every circle with every triangle, showing that the boxes did have the same size! This idea is what we will use.

Definition 5.1. *A set A is **equinumerous** to a set B, denoted $A \approx B$, iff there is a injective function A onto B. Such a function is called an **one-to-one correspondence** between A and B.*

Example 3. *Using a diagonalization argument, one can show that $\omega \approx \omega \times \omega$. We define a map $J : \omega \times \omega \to \omega$ by*

$$J(m, n) = [1 + 2 + \cdots + (m + n)] + m = \frac{1}{2}[(m + n)^2 + 3m + n]$$

This map is indeed a one-to-one correspondence between these sets. Geometrically, this amounts to drawing diagonal lines on $\omega \times \omega$, where each line contains the ordered pairs whose coordinates sum to the same value. Consider the point $\langle k, m\rangle$ so that $k + m = n$. Note that to get to a diagonal we must count all diagonals below it. Convince yourself that the dialgonal whose elements sum to p has $p + 1$ elements. Hence for $\langle k, m\rangle$ we must count at least

$$(0 + 1) + (1 + 1) + \cdots + ((m - 1) + 1) = \sum_{i=1}^{m} i = \frac{m(m + 1)}{2}$$

The number of elements in these diagonals corresponds to the value

$$\frac{(k+m)(k+m+1)}{2}$$

and since $\langle k, m\rangle$ is the $k+1$th element of the diagonal, we add k, giving us

$$\frac{(k+m)(k+m+1)}{2} + k = \frac{1}{2}((k+m)^2 + 3k + n)$$

To show then that this map is injective, let $J(\langle k_0, m_0\rangle) = J(\langle k_1, m_1\rangle)$, where $k_0 + m_0 = n_0$ and $k_1 + m_1 = n_1$. Assume that $n_0 \neq n_1$ and WLOG let $n_0 > n_1$. The map

$$\tau(x) = \frac{x(x+1)}{2}$$

is strictly increasing on the natural numbers. Thus,

$$\begin{aligned}
\tau(n_0) > \tau(n_1) &\implies \tau(n_0) \geq \tau(n_1 + 1) \\
&\implies \tau(n_0) \geq \frac{(n_1+1)(n_1+2)}{2} \\
&\implies \tau(n_0) \geq \frac{n_1(n_1+1)}{2} + n_1 + 1 \\
&\implies \tau(n_0) \geq \tau(n_1) + n_1 + 1 \\
&\implies \tau(n_0) \geq \tau(n_1) + k_1
\end{aligned}$$

But from this, we get that

$$J(\langle k_0, m_0\rangle) = \tau(n_0) + k_0 \geq \tau(n_0) \geq \tau(n_1) + k_1 = J(\langle k_1, m_1\rangle)$$

a contradiction.
Similarly, one can show that $\omega \approx \mathbb{Q}$.

Example 4. *The open unit interval $(0, 1)$ is equinumerous to $\mathbb{R}$. Bending $(0, 1)$ into a semicircle P, we can form a one-to-one correspondence with $\mathbb{R}$ by projecting each point of P onto the real line.*

Example 5. *For any set A, the power set $\mathcal{P}A$ is equinumerous to the set of all functions from A to 2. Indeed, for any $B \subseteq A$, we define $f_B : A \to 2$ by*

$$f_B(x) = \begin{cases} 1 & x \in B \\ 0 & x \in A \setminus B \end{cases}$$

Similarly, any function g from A to 2 can be paired with a subset of A, by taking $\{x \in A : g(x) = 1\}$.

One can define an equivalence relation E by saying two sets A, B are equivalent if they're equinumerous.

Theorem 5.2. *For any sets A, B, C,*

(i) $A \approx A$

(ii) *If* $A \approx B$ *then* $B \approx A$

(iii) *If* $A \approx B$ *and* $B \approx C$*, then* $A \approx C$

One might think, in light of the above examples, that any two infinte sets are equinumerous. This is not true; some infinite sets are much larger than others:

Theorem 5.3 (Cantor's Theorem (1873))**.** *(i)* ω *is not equinumerous to* $\mathbb{R}$

(ii) No set is equinumerous to its power set

Proof. (i) Let $f : \omega \to \mathbb{R}$. We claim there is a real z not in the range of f. Suppose that no such z exists. Then we can enumerate all the real numbers as outputs of f like so:

$$\begin{aligned} f(0) &= 240.013\ldots, \\ f(1) &= -7.456\ldots, \\ f(2) &= 1.14141\ldots, \\ &\vdots \end{aligned}$$

We now define z as follows: The integer component is 0, while the $(n+1)$th decimal place is 7, unless the $(n+1)$th decimal place of $f(n)$ is 7, in which case we set the $(n+1)$th decimal place of z to 6. Notice that for every $f(n)$, z differs from $f(n)$ in at least one decimal place, so they're not the same number. Thus, z cannot be in the range of f.

(ii) Assume there is a one-to-one correspondence from a set A onto its power set $\mathcal{P}A$. It follows that for all $B \subseteq A$, there is an $a \in A$ such that $g(a) = B$. Let

$$\mathcal{C} = \{x \in A : x \notin g(x)\}$$

Then for some $c \in A$, $g(c) = \mathcal{C}$. But, notice that for c,

$$c \in g(c) \iff c \notin \mathcal{C}$$

which is a contradiction.

□

5.2 FINITE SETS

While we've used the words "finite" and "infinite" informally before, we now define them precisely.

Definition 5.4. *A set is* ***finite*** *if and only if it is equinumerous to some natural number. Otherwise, we call the set* ***infinite****.*

In constructing the naturals, we defined each number as a set containing all smaller natural numbers; this is crucial for our definition.

We must check that each finite set is equinumerous to a unique number n. Then we can use n as the number of elements in S. To do this, we require a theorem, which tells us that if we have n objects and few than n places to put them, one place will have have more than one object. You know this by a famous name:

Theorem 5.5 (The Pigeonhole Principle)**.** *No natural number is equinumerous to a proper subset of itself.*

Note that a set is a proper subset if it is a subset and not equal to the superceding set.

Proof. Let f be one-to-one from n to n. We will show that the range of n is n itself, which proves the claim.

We use induction. To this end, define

$$T = \{n \in \omega : \text{ any one-to-one function from } n \text{ into } n \text{ has range } n\}$$

and claim $T = \omega$. Clearly $0 \in T$, as the only function from 0 to 0 is ω, whose range is 0. Now suppose that $k \in T$ and f is one-to-one from k^+ to k^+.

Notice that $f \restriction k$ maps k one-to-one into k^+. We thus have two cases.

1. Suppose k is closed under f. Then $f \restriction k$ maps k into k. Because $k \in T$ we know then that the range of this restriction is k. As f is one-to-one the only value left for $f(k)$ is k. Thus,
$$\operatorname{ran} f = k \cup \{k\} = k^+$$
2. Suppose that $f(p) = k$ for some p smaller than k. We can swap two values of f as follows: define $\hat{f}$ by
$$\hat{f}(p) = f(k), \quad \hat{f}(k) = f(p)$$
and $\hat{f}(x) = f(x)$ for all other $x \in k^+$. $\hat{f}$ maps k^+ one-to-one into k^+, and k is closed under $\hat{f}$, so the first case applies and $\operatorname{ran} \hat{f} = k^+$. But $\operatorname{ran} \hat{f} = \operatorname{ran} f$.

In either case, the range of f is k^+, so $k^+ \in T$. We conclude that $T = \omega$, as required. □

For a finite set A, the unique n for which $A \approx n$ is called the **cardinal number** of A, and is denoted $\operatorname{card}(A)$.

Example 6. *If a, b, c, d are distinct, then* $\operatorname{card}(\{a, b, c, d\}) = 4$.

For infinite sets, it's a bit more complicated so we postpone it until a later lecture.

We define cardinal numbers to be the anything that is the cardinal number of some set A. All natural numbers are cardinal numbers, since $\operatorname{card}(n) = n$. $\operatorname{card}(\omega)$ is not a natural number, since ω is not equinumerous to any natural number. We will not reveal exactly what the cardinality of ω is until later, but we will give it a name:

$$\operatorname{card}(\omega) = \aleph_0$$

5.3 CARDINAL ARITHMETIC

Addition, multiplication, and exponentiation are useful for finite cardinals, and are also useful for arbitrary cardinals as well. We need to extend these operations to the infinite cardinals.

While our previous derivations for operations on ω won't work, we can find a suitable derivation. Remember in elementary school when you were asked to do $2 + 3$. You did not use the Recursion Theorem to prove this. Instead, you got two sets, one of size 2, and the other of size 3, and showed that their union is a set of size 5; the sets could be fingers, or apples, or pencils. The exact same idea works for cardinal numbers:

Definition 5.6. *Let κ, λ be any cardinal numbers. Then*

(i) $\kappa + \lambda = \mathrm{card}(K \cup L)$, *where K, L are disjoint sets of cardinality κ and λ respectively.*

(ii) $\kappa \cdot \lambda = \mathrm{card}(K \times L)$, *where K, L are any sets of cardinality κ and λ respectively.*

(iii) $\kappa^\lambda = \mathrm{card}(L^K)$, *where K, L are any sets of cardinality κ and λ respectively.*

5.4 ORDERING CARDINAL NUMBERS