Notes on ordinals and cardinals

Reed Solomon

1 Background Terminology

We will use the following notation for the common number systems:

$$\begin{gathered}\mathbb{N} = \{0, 1, 2, \ldots\} = \text{ the natural numbers} \\ \mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\} = \text{the integers} \\ \mathbb{Q} = \{m/n \mid m, n \in \mathbb{Z} \wedge n \neq 0\} = \text{ the rational numbers} \\ \mathbb{R} = \text{ the real numbers}\end{gathered}$$

Definition 1.1. Let A and B be sets. The *Cartesian product* of A and B is

$$A \times B = \{\langle a, b\rangle \mid a \in A \text{ and } b \in B\}.$$

We write A^2 for $A \times A$. In general, A^n is the set of n-tuples of the form $\langle a_0, a_1, \ldots, a_{n-1}\rangle$ where each $a_i \in A$.

Definition 1.2. Let A be a set and $n \in \mathbb{N}$ with $n \geq 1$. An *n-ary relation R on A* is a subset $R \subseteq A^n$. Given a tuple $\overline{a} = \langle a_0, \ldots, a_{n-1}\rangle \in A^n$, we say *$R$ holds of $\overline{a}$* if $\overline{a} \in R$, and otherwise we say *R does not hold of $\overline{a}$*.

If $n = 1$, we call R a *unary relation*; if $n = 2$, we call R a *binary relation*; and if $n = 3$, we call R a *ternary relation*. In these notes, we will mostly consider binary relations, typically specifying some type of ordering relation.

Definition 1.3. Let R be a binary relation on a set A.

- R is *reflexive* if for all $a \in A$, $\langle a, a\rangle \in R$.
- R is *irreflexive* if for all $a \in A$, $\langle a, a\rangle \notin R$.
- R is *symmetric* if for all $a, b \in A$, $\langle a, b\rangle \in R$ implies $\langle b, a\rangle \in R$.
- R is *antisymmetric* if for all $a, b \in A$, $\langle a, b\rangle \in R$ and $\langle b, a\rangle \in R$ implies that $a = b$.
- R is *transitive* if for all $a, b, c \in A$, if $\langle a, b\rangle \in R$ and $\langle b, c\rangle \in R$, then $\langle a, c\rangle \in R$.
- R is *total* if for all $a, b \in A$, either $\langle a, b\rangle \in R$ or $\langle b, a\rangle \in R$.

- R satisfies *trichotomy* if for all $a, b \in A$, exactly one of $a = b$, $\langle a, b\rangle \in R$ or $\langle b, a\rangle \in R$ holds.

For the most part, our examples come from the following classes of algebraic structures.

Definition 1.4. An *equivalence relation* is a pair $(E, \sim)$ such that E is a set and $\sim$ is a binary relation on E which is reflexive, transitive and symmetric.

Example 1.5. Let A and B be sets and let $f : A \to B$ be a function between them. Define a binary relation $\sim$ on A by $a \sim b$ if and only if $f(a) = f(b)$. The relation $\sim$ is reflexive because $f(a) = f(a)$ (i.e. $a \sim a$) for all $a \in A$. It is symmetric because for any $a, b \in A$, if $f(a) = f(b)$ (i.e. $a \sim b$), then $f(b) = f(a)$ (i.e. $b \sim a$). It is transitive because for any $a, b, c \in A$, if $f(a) = f(b)$ (i.e. $a \sim b$) and $f(b) = f(c)$ (i.e. $b \sim c$), then $f(a) = f(c)$ (i.e. $a \sim c$).

Definition 1.6. A *partial order* is a pair $(P, \leq_P)$ such that P is a set and $\leq_P$ is a binary relation on P which is reflexive, antisymmetric and transitive.

Example 1.7. Consider the set F of all finite binary strings. That is,

$$F = \{\lambda, \langle 0\rangle, \langle 1\rangle, \langle 0, 0\rangle, \langle 0, 1\rangle, \langle 1, 0\rangle, \langle 1, 1\rangle, \ldots\}$$

where λ denotes the empty string. We typically denote arbitrary finite binary strings by σ, τ and δ. Define a binary relation $\preceq$ on F by $\sigma \preceq \tau$ if and only if σ is an initial segment of τ. (This initial segment does not need to be proper.) For example,

$$\langle 1, 0\rangle \preceq \langle 1, 0, 1, 1\rangle \text{ and } \langle 0, 0, 0\rangle \preceq \langle 0, 0, 0\rangle$$

The relation $\preceq$ is a partial order on F. It is reflexive because each string is an initial segment of itself (i.e. $\sigma \preceq \sigma$). It is antisymmetric because if σ is an initial segment of τ (i.e. $\sigma \preceq \tau$) and τ is an initial segment of σ (i.e. $\tau \preceq \sigma$), then $\sigma = \tau$. It is transitive because if σ is an initial segment of τ and τ is an initial segment of δ, then σ is an initial segment of δ.

Later when we have introduced more set theoretic notation, we will denote F by either $\{0, 1\}^{<\omega}$ or $2^{<\omega}$. It is also common in computer science or in the study of formal languages to denote F by $\{0, 1\}^*$.

Definition 1.8. A *linear order* is a pair $(L, \leq_L)$ such that L is a set and $\leq_L$ is a binary relation on L which is reflexive, antisymmetric, transitive and total. (Thus, a linear order is a partial order which is total.)

Example 1.9. The standard orders on the number usual number systems are all linear orders. For example, $(\mathbb{N}, \leq_{\mathbb{N}})$, $(\mathbb{Z}, \leq_{\mathbb{Z}})$, $(\mathbb{Q}, \leq_{\mathbb{Q}})$ and $(\mathbb{R}, \leq_{\mathbb{R}})$ are all linear orders.

In Definition 1.8, we axiomatized a binary relation representing "less than or equal to". We can also describe linear orders by axiomatizing the "less than and not equal to" relation. To distinguish these two notions, we refer to the structures axiomatized by the "less than and not equal to" relation as strict linear orders.

Definition 1.10. A *strict linear order* is a pair $(L, <_L)$ such that L is a set and $<_L$ is a binary relation which is irreflexive, transitive and satisfies trichotomy.

Example 1.11. $(\mathbb{N}, <_{\mathbb{N}})$, $(\mathbb{Z}, <_{\mathbb{Z}})$, $(\mathbb{Q}, <_{\mathbb{Q}})$ and $(\mathbb{R}, <_{\mathbb{R}})$ are all strict linear orders.

One can pass between linear orders and strict linear orders in a natural way. If you have not seen these definitions before, it is worth working through the following two exercises.

Exercise 1.12. Let $(L, <_L)$ be a strict linear order. Define the binary relation $\leq_L$ on L by $x \leq_L y$ if and only if $x <_L y$ or $x = y$. Show that $(L, \leq_L)$ is a linear order.

Exercise 1.13. Let $(L, \leq_L)$ be a linear order. Define the binary relation $<_L$ on L by $x <_L y$ if and only if $x \leq_L y$ and $x \neq y$. Show that $(L, <_L)$ is a strict linear order.

We will also use the following terminology for functions.

Definition 1.14. Let A and B be sets and let $f : A \to B$ be a function between them. The *range of* f is the set

$$\text{range}(f) = \{b \in B \mid \exists a \in A\,(f(a) = b)\}$$

We say f is *one-to-one* (or *injective* or is *an injection*) if $f(a_1) \neq f(a_2)$ whenever $a_1 \neq a_2 \in A$. In other words, if $b \in \text{range}(f)$, then there is a unique element $a \in A$ such that $f(a) = b$.

We say f is *onto* B (or is *surjective* or is *a surjection*) if $\text{range}(f) = B$; that is, if for every $b \in B$, there is an $a \in A$ such that $f(a) = b$. We say f is a *bijection* if f is both injective and surjective.

If $f : A \to B$ is one-to-one, then we can define the inverse $f^{-1} : \text{range}(f) \to A$ of f by $f^{-1}(b) = a$ where a is the unique element of A such that $f(a) = b$. In particular, if $f : A \to B$ is a bijection, then $f^{-1} : B \to A$ is also a bijection.

Definition 1.15. Let $f : A \to B$ be a function and let $C \subseteq A$ be a subset of the domain.

We define $f \restriction C : C \to B$, called *the restriction of* f *to* C, by $f \restriction C\,(x) = f(x)$ for all $x \in C$. That is, we leave the function the same except we restrict its domain to C.

We define the *range of* f *on* C, denoted, $f[C]$ by $f[C] = \{b \in B \mid \exists c \in C\,(f(c) = b)\}$. That is, $f[C] = \text{range}(f \restriction C)$. Note that $f[A] = \text{range}(f)$.

Definition 1.16. Let A be a set. The *power set of* A, denoted $\mathcal{P}(A)$, is defined by

$$\mathcal{P}(A) = \{B \mid B \subseteq A\}$$

Example 1.17. If $A = \{a, b, c\}$, then

$$\mathcal{P}(A) = \{\,\emptyset,\, \{a\},\, \{b\},\, \{c\},\, \{a, b\},\, \{a, c\},\, \{b, c\},\, \{a, b, c\}\,\}.$$

Definition 1.18. If A and B are sets, B^A denotes the set of all functions from A into B.

Example 1.19. $\mathbb{R}^{\mathbb{N}}$ denotes the set of all functions from $\mathbb{N}$ to $\mathbb{R}$. That is,

$$\mathbb{R}^{\mathbb{N}} = \{f \mid f : \mathbb{N} \to \mathbb{R}\}.$$

Notice that this set is not the same as $\mathbb{N}^{\mathbb{R}}$ which is all the function $f : \mathbb{R} \to \mathbb{N}$.

Example 1.20. It is worth thinking about the set A^B when either A or B is the empty set. Suppose $B = \emptyset$ and consider $A^\emptyset$. $A^\emptyset$ is the set of all functions $f : \emptyset \to A$. There is exactly one such function, namely the empty function. So, $A^\emptyset = \{\emptyset\}$ for any set A (even if $A = \emptyset$).

Consider the case when $A = \emptyset$ and $B \neq \emptyset$. In this case, $\emptyset^B$ is the set of all functions $f : B \to \emptyset$. Since B is not empty, there is some $b \in B$. Such a function f must be defined on b, i.e. $f(b)$ must be defined. But the range of f has to be contained in $\emptyset$, so there are no possible values for $f(b)$. Therefore, there cannot be such a function f and so $\emptyset^B = \emptyset$ when $B \neq \emptyset$.

For reasons that will become clear later, mathematicians sometimes use a natural number n to denote the set of natural numbers strictly less than n. For example, $0 = \emptyset$, $1 = \{0\}$, $2 = \{0, 1\}$ and so on. You have to rely on context to tell you what is meant, but the main place that numbers are used to stand for sets is in the B^A notation. For example, $2^{\mathbb{N}}$ denotes the set of all functions $f : \mathbb{N} \to \{0, 1\}$.

At various places, it will be useful to use the Axiom of Choice. We will use it in various forms throughout the course, but we begin with the most standard form of this axiom.

Axiom 1.21 (Axiom of Choice)**.** *Let $\mathcal{F} = \{U_a \mid a \in A\}$ be a family of nonempty sets indexed by a nonempty set A. There is a function $f : A \to \bigcup_{a \in A} U_a$ such that for every $a \in A$, $f(a) \in U_a$.*

2 Comparing sizes of sets

Definition 2.1. We say that two sets A and B *have the same size* or *have the same cardinality*, and write $|A| = |B|$, if there is a bijection $f : A \to B$.

Exercise 2.2. Prove that the relation of having the same size is an equivalence relation. That is, prove that this relation is reflexive ($|A| = |A|$), symmetric (if $|A| = |B|$, then $|B| = |A|$) and transitive (if $|A| = |B|$ and $|B| = |C|$, then $|A| = |C|$).

Example 2.3. $|\{0, 1\}| = |\{3, 4\}|$ because there is a bijection $f : \{0, 1\} \to \{3, 4\}$ given by $f(0) = 3$ and $f(1) = 4$. Notice that this bijection is not unique.

Example 2.4. Let $\text{Even} = \{0, 2, 4, \ldots\}$ and $\text{Odd} = \{1, 3, 5, \ldots\}$. Then $|\mathbb{N}| = |\text{Even}|$ by $f(x) = 2x$, $|\mathbb{N}| = |\text{Odd}|$ by $g(x) = 2x + 1$ and $|\text{Even}| = |\text{Odd}|$ by $h(x) = x + 1$. Notice that having shown $|\mathbb{N}| = |\text{Even}|$ and $|\mathbb{N}| = |\text{Odd}|$, we could immediately conclude $|\text{Even}| = |\text{Odd}|$ because the relation of having the same size is an equivalence relation.

Example 2.5. Let $\mathbb{R}^+ = \{r \in \mathbb{R} \mid r > 0\}$. $|\mathbb{R}| = |\mathbb{R}^+|$ because of the bijection $f(x) = e^x$.

Example 2.6. Let $a, b \in \mathbb{R}$ with $a < b$ and let (a, b) denote the open interval

$$(a, b) = \{r \in \mathbb{R} \mid a < r < b\}.$$

We have $|(a, b)| = |(0, 1)|$ because of the function $f(x) = (x - a)/(b - a)$. Therefore, each nontrivial open interval in the real line has the same size. Moreover, $|\mathbb{R}| = |(-\pi/2, \pi/2)|$ because of the function $g(x) = \tan^{-1}(x)$. Because the relation of having the same size is an equivalence relation, $|\mathbb{R}| = |(a, b)|$ for any $a < b$ in $\mathbb{R}$.

Definition 2.7. A set B is called *countable* if $|B| = |\mathbb{N}|$ and is said to have *size continuum* if $|B| = |\mathbb{R}|$.

Example 2.8. The following sets are all countable: Even, Odd, $\mathbb{Z}$ and $\mathbb{Q}$.

Example 2.9. The set $\mathbb{R}^+$ has size continuum as does any nontrivial open interval (a, b) in $\mathbb{R}$.

Exercise 2.10. Show that $|\mathcal{P}(\mathbb{N})| = |2^{\mathbb{N}}|$. Recall that $2^{\mathbb{N}}$ really means $\{0, 1\}^{\mathbb{N}}$, i.e. the set of functions of the form $f : \mathbb{N} \to \{0, 1\}$.

For the examples so far, we defined our bijections explicitly. This method is unnecessarily restrictive and it is quite useful to develop some basic tools that allow one to conclude that sets have the same size without giving an explicit map. Here is one simple version of such a tool which we will give in a more general form later.

Theorem 2.11. *Let A be an infinite set. If there is a one-to-one function $f : A \to \mathbb{N}$, then A is countable.*

Proof. Let A be infinite and assume that $f : A \to \mathbb{N}$ is one-to-one. We can write

$$\text{range}(f) = \{n_0, n_1, n_2, \ldots\} \text{ where } n_0 <_{\mathbb{N}} n_1 <_{\mathbb{N}} \cdots .$$

That is, we write range(f) in increasing order. Since f is one-to-one, it has an inverse $f^{-1} : \text{range}(f) \to A$. Notice that f^{-1} is onto A. That is, for every $a \in A$, there is an element $n_k \in \text{range}(f)$ such that $f^{-1}(n_k) = a$, namely let $n_k = f(a)$.

We define a bijection $g : \mathbb{N} \to A$ by setting $g(k) = f^{-1}(n_k)$. To see that g is a bijection, we have to check that it is one-to-one and onto. Since f^{-1} is onto, it follows that g is onto.

To see that f^{-1} is one-to-one, fix $k \neq \ell \in \mathbb{N}$ and assume for a contradiction that $g(k) = g(\ell)$. By the definition of g, this means $f^{-1}(n_k) = f^{-1}(n_\ell)$. But then $f(f^{-1}(n_k)) = f(f^{-1}(n_\ell))$ and hence $n_k = n_\ell$. □

Let me give a couple of applications of Theorem 2.11.

Theorem 2.12. *If A and B are countable, then so are $A \cup B$ and $A \times B$.*

Proof. First, consider the case of $A \cup B$. Since $A \cup B$ is infinite, it suffices to give a one-to-one function $f : A \cup B \to \mathbb{N}$. Because A and B are countable, we can fix bijections $g : A \to \mathbb{N}$ and $h : B \to \mathbb{N}$. Define f by

$$f(x) = \begin{cases} 2g(x) & \text{if } x \in A \\ 2h(x) + 1 & \text{otherwise} \end{cases}$$

That is, f maps the elements of A into the set Even $\subseteq \mathbb{N}$ by doubling the value of $g(x)$. If $x \notin A$, then f uses the function h suitably modified to land in the set Odd $\subseteq \mathbb{N}$. In particular, f is a map from $A \cup B$ into $\mathbb{N}$.

To see that f is one-to-one, consider a pair of elements $x \neq y \in A \cup B$. If $x \in A$ and $y \notin A$, then $f(x)$ is even and $f(y)$ is odd, so $f(x) \neq f(y)$. (Similarly, if $x \notin A$ and $y \in A$,

then $f(x) \neq f(y)$.) Therefore, assume that x and y come from the same set. If $x, y \in A$, then $f(x) = 2g(x)$ and $f(y) = 2g(y)$. Since $g(x) \neq g(y)$ (because g is one-to-one), we have $2g(x) \neq 2g(y)$ and hence $f(x) \neq f(y)$. The argument when $x, y \in B$ is similar.

The function f is not necessarily onto $\mathbb{N}$. It is well worth coming up with an example in which the function f given here is not onto.

Second, consider the case of $A \times B$. As above, since $A \times B$ is infinite, it suffices to give a one-to-one function $f : A \times B \to \mathbb{N}$. As above, fix bijections $g : A \to \mathbb{N}$ and $h : B \to \mathbb{N}$. We define $f(\langle a, b\rangle) = 2^{g(a)}3^{h(b)}$. Notice that f maps into $\mathbb{N}$ as required.

To show that f is one-to-one, suppose $f(\langle a, b\rangle) = f(\langle c, d\rangle)$ and we show that $\langle a, b\rangle = \langle c, d\rangle$, i.e. that $a = c$ and $b = d$. Because $f(\langle a, b\rangle) = f(\langle c, d\rangle)$, we have $2^{g(a)}3^{h(b)} = 2^{g(c)}3^{h(d)}$. Because natural numbers have unique prime factorizations, $g(a) = g(c)$ and $h(b) = h(d)$. But, g and h are one-to-one, so $a = c$ and $b = d$ as required. □

The following theorem can be proved in many ways, but one way is to mimic the proof above for the case of $A \times B$. I will leave it to you to think about.

Theorem 2.13. *A countable union of countable sets is countable.*

Theorem 2.14. *For each $k \geq 1$, the set $\mathbb{N}^k$ is countable.*

Proof. We proceed by induction on k. The base case is when $k = 1$ and it says that $\mathbb{N}$ is countable which is clearly true.

For the induction, we assume that $\mathbb{N}^k$ is countable and we show that $\mathbb{N}^{k+1}$ is countable. By Theorem 2.12, $\mathbb{N}^k \times \mathbb{N}$ is countable. However, $|\mathbb{N}^k \times \mathbb{N}| = |\mathbb{N}^{k+1}|$ by the map that sends $\langle\langle n_1, n_2, \ldots, n_k\rangle, n_{k+1}\rangle$ to $\langle n_1, n_2, \ldots, n_k, n_{k+1}\rangle$. □

Corollary 2.15. *If A is countable, then for each $k \geq 1$, A^k is countable.*

Proof. Fix a bijection $f : A \to \mathbb{N}$. For any $k > 1$, we can define a bijection $g_k : A^k \to \mathbb{N}^k$ by

$$g_k(\langle a_0, a_1, \ldots, a_{k-1}\rangle) = \langle f(a_0), f(a_1), \ldots, f(a_{k-1})\rangle.$$

Therefore $|A^k| = |\mathbb{N}^k|$ and hence A^k is countable. □

Combining Theorems 2.13 and 2.14, we have that

$$\mathbb{N}^{<\omega} = \bigcup_{k \geq 1} \mathbb{N}^k$$

is countable. Notice that $\mathbb{N}^{<\omega}$ is the set of all finite sequences of natural numbers. There is nothing important about the fact that the base set is $\mathbb{N}$ rather than some other countable set. Therefore, we have that if A is a nonempty finite or countable set, then the set of all finite sequences of elements of A, denoted $A^{<\omega}$ or A^*, is countable.

Hopefully these examples will have illustrated the importance of the existence of a one-to-one function $f : A \to B$ between two sets. We give a notation for this concept in the next definition.

Definition 2.16. We say that the set A is *at most as big as the set* B, and write $|A| \leq |B|$ if there is a one-to-one function $f : A \to B$.

Note that this includes the case when $A = \emptyset$ for which we regard the empty map as a one-to-one function from A into any set B. In other words, $|\emptyset| \leq |B|$ for every set B.

Exercise 2.17. Show that if $|A| \leq |B|$ and $|B| \leq |C|$, then $|A| \leq |C|$.

Example 2.18. If $A \subseteq B$, then $|A| \leq |B|$ by the inclusion map $f : A \to B$ given by $f(a) = a$.

Example 2.19. If A is an infinite set, then $|\mathbb{N}| \leq |A|$. We need to define a one-to-one function $f : \mathbb{N} \to A$. We define this function by recursion on $\mathbb{N}$. Defining functions by recursion is analogous to induction proofs. That is, for the base case, we specify $f(0)$. For the induction case, we assume that we have specified $f(0), \ldots, f(n)$ and we specify $f(n+1)$ using our knowledge about the values $f(0), \ldots, f(n)$.

For the base case, we set $f(0) = a_0$ for some arbitrary element $a_0 \in A$. For the induction case, assume that we have defined the values of $f(0), \ldots, f(n)$. We need to specify $f(n+1)$ and we would like to make sure that $f(n+1)$ is not equal to any of $f(0), \ldots, f(n)$ so that f has a chance to be one-to-one. To specify $f(n+1)$, notice that

$$A \setminus \{f(0), f(1), \ldots, f(n)\}$$

is nonempty because A is infinite and we are only removing finitely many elements. Therefore, we can set $f(n+1)$ to be any arbitrary element of $A \setminus \{f(0), \ldots, f(n)\}$, i.e. we pick an element $a_{n+1} \in A \setminus \{f(0), \ldots, f(n)\}$ and set $f(n+1) = a_{n+1}$.

This completes the definition of f. To see that f is one-to-one, fix $n \neq m \in \mathbb{N}$. We need to show that $f(n) \neq f(m)$. We can assume without loss of generality that $n <_{\mathbb{N}} m$. Consider the induction step in the definition of f when we define $f(m)$. The value of $f(m)$ is chosen from $A \setminus \{f(0), \ldots, f(n), \ldots, f(m-1)\}$. In particular, $f(m)$ is specifically chosen so that $f(m) \neq f(n)$.

We have considered the case when we have a one-to-one map $f : A \to B$. The next lemma shows that we can capture the same connection in a dual manner by considering onto maps $g : B \to A$.

Lemma 2.20. *For any nonempty sets A and B, $|A| \leq |B|$ if and only if there is an onto function $g : B \to A$.*

Proof. We prove the two implications separately. For the first direction, assume that $|A| \leq |B|$. By Definition 2.16, we can fix an one-to-one $f : A \to B$. Because A is nonempty, we can fix an element $a_0 \in A$. We define the map $g : B \to A$ as follows.

$$g(b) = \begin{cases} f^{-1}(b) & \text{if } b \in \text{range}(f) \\ a_0 & \text{otherwise} \end{cases}$$

Note that if b is in the range of f, then $f^{-1}(b)$ is well defined because f is injective. To see that g is surjective, fix an arbitrary element $a \in A$. Let $b = f(a)$. Then, since $b \in \text{range}(f)$, we have $g(b) = f^{-1}(b) = a$. Therefore, g is onto A as required.

For the second direction, assume that $g : B \to A$ is onto. We will use the Axiom of Choice to define a one-to-one $f : A \to B$. For each $a \in A$, let $U_a = \{b \in B \mid g(b) = a\}$ be the set of elements of B which map to a under g. Because g is onto, the set U_a is nonempty for each $a \in A$. Therefore, we have the following family of nonempty sets $\mathcal{F} = \{U_a \mid a \in A\}$ indexed by A. By the Axiom of Choice, there is a function $f : A \to \bigcup_{a \in A} U_a$ such that $f(a) \in U_a$ for each $a \in A$.

We check that f has the desired properties. First, we show that f maps A into B. Since $U_a \subseteq B$ for each $a \in A$, we have $\bigcup_{a \in A} U_a \subseteq B$ and hence $f : A \to B$.

Next, we show that f is one-to-one. Fix $u \neq v \in A$ and we show that $f(u) \neq f(v)$. We claim that $U_u \cap U_v = \emptyset$. To see why, fix an arbitrary $b \in U_u$. By the definition of U_u, $g(b) = u$ and hence $g(b) \neq v$. Therefore, by the definition of U_v, $b \notin U_v$ and hence $U_u \cap U_v = \emptyset$ as claimed. Since $f(u) \in U_u$, $f(v) \in U_v$ and $U_u \cap U_v = \emptyset$, it follows that $f(u) \neq f(v)$ as required. □

Before proceeding, we should notice a curious phenomenon. Suppose that we have a set A such that $|\mathbb{N}| \leq |A|$ and $|A| \leq |\mathbb{N}|$. Because $|\mathbb{N}| \leq |A|$, we know that A is infinite. (Why?) But, if A is infinite and $|A| \leq |\mathbb{N}|$, then A is countable and hence $|A| = |\mathbb{N}|$. In other words, $|\mathbb{N}| \leq |A|$ and $|A| \leq |\mathbb{N}|$ implies that $|A| = |\mathbb{N}|$. The next theorem says that this property holds for all sets and not just for countable sets. You will prove it in the homework.

Theorem 2.21 (Schroeder-Bernstein Theorem)**.** *For any sets A and B, $|A| = |B|$ if and only if $|A| \leq |B|$ and $|B| \leq |A|$.*

The Schroeder-Bernstein Theorem is very useful for determining cardinalities without giving explicit bijections.

Example 2.22. $\mathbb{Q}$ is countable. Since $\mathbb{N} \subseteq \mathbb{Q}$, we have $|\mathbb{N}| \leq |\mathbb{Q}|$. To prove the other inequality, we show that $|\mathbb{Q}| \leq |\mathbb{N} \times \mathbb{N}|$. Since $|\mathbb{N} \times \mathbb{N}| = |\mathbb{N}|$, we obtain $|\mathbb{Q}| \leq |\mathbb{N}|$ and hence $|\mathbb{Q}| = |\mathbb{N}|$ as required.

To show $|\mathbb{Q}| \leq |\mathbb{N} \times \mathbb{N}|$, we define a one-to-one function $f : \mathbb{Q} \to \mathbb{N} \times \mathbb{N}$. Our definition of $f(q)$ splits into cases depending on whether $q \geq 0$ or $q < 0$. If $q \geq 0$, then we write $q = m/n$ in reduced form with $m, n \in \mathbb{N}$. (Reduced form means that n and m have no divisors in common.) Define $f(q) = \langle 2m, n \rangle$. If $q < 0$, then we write $q = -m/n$ in reduced form with $m, n \in \mathbb{N}$. Define $f(q) = \langle 2m + 1, n \rangle$.

We see that f is one-to-one, fix $q \neq q' \in \mathbb{Q}$. If $q \geq 0$ and $q' < 0$, then $f(q) \neq f(q')$ because the first component of $f(q)$ is even and the first component of $f(q')$ is odd. Similarly, if $q = 0$ and $q' \neq 0$, then $f(q) \neq f(q')$ because the first component of $f(q)$ is 0 and the first component of $f(q')$ is not 0. Therefore, assume without loss of generality that q and q' are either both positive or both negative. We consider the case when $q = m/n$ and $q' = m'/n'$ are both positive as the case when both are negative is similar. When both are positive, either $m \neq m'$ or $n \neq n'$. If $m \neq m'$, then $2m \neq 2m'$ and $f(q) \neq f(q')$ because they differ on the first component. If $n \neq n'$, then $f(q) \neq f(q')$ because they differ on the second component.

Example 2.23. For any $a <_{\mathbb{R}} b$ in $\mathbb{R}$, we have $|[a, b]| = |\mathbb{R}|$. Since $[a, b] \subseteq \mathbb{R}$, we have $|[a, b]| \leq \mathbb{R}$. To show that other inequality, notice that $|\mathbb{R}| = |(a, b)|$ by an earlier example and $|(a, b)| \leq |[a, b]|$ because $(a, b) \subseteq [a, b]$. Therefore, $|\mathbb{R}| \leq |[a, b]|$ are required.

Example 2.24. $|\mathbb{R}| = |2^{\mathbb{N}}| = |\mathcal{P}(\mathbb{N})|$. The second equality was an exercise above. To see the first equality, we prove the two inequalities.

First, consider $|2^{\mathbb{N}}| \leq |\mathbb{R}|$. We define a one-to-one function $\alpha : 2^{\mathbb{N}} \to \mathbb{R}$. Recall that each element of $2^{\mathbb{N}}$ is a function $f : \mathbb{N} \to \{0, 1\}$. For such a function f, we define

$$\alpha(f) = \sum_{n=0}^{\infty} \frac{2f(n)}{3^{n+1}}$$

This map sends f into Cantor's Middle Thirds set. You can check that it is one-to-one. (If $f \neq g$, consider the least n such that $f(n) \neq g(n)$. Assume $f(n) = 0$ and $g(n) = 1$. By considering the tails of these infinite sums, show that even if $f(k) = 1$ for all $k > n$ and $g(k) = 0$ for all $k > n$, you still have $\alpha(f) < \alpha(g)$.)

Second, consider $|\mathbb{R}| \leq |2^{\mathbb{N}}|$. Since $|\mathbb{N}| = |\mathbb{Q}|$, we have $|2^{\mathbb{N}}| = |2^{\mathbb{Q}}| = |\mathcal{P}(\mathbb{Q})|$. Therefore, it suffices to show that $|\mathbb{R}| \leq |\mathcal{P}(\mathbb{Q})|$. Define $f : \mathbb{R} \to \mathcal{P}(\mathbb{Q})$ by $f(r) = \{q \in \mathbb{Q} \mid q < r\}$. That is, f maps r to the Dedkind cut determined by r.

To see that f is one-to-one, fix $r \neq r' \in \mathbb{R}$ and assume that $r < r'$. Because the rational numbers are dense in $\mathbb{R}$, there is a $q \in (r, r')$. By definition, $q \in f(r')$ but $q \notin f(r)$, so $f(r') \neq f(r)$.

Hopefully, one of the lessons that emerges from these applications of the Schroeder-Bernstein Theorem is that it is important to pick the right inequalities to show. That is, it is often useful to pick the correct sets to work with and this often involves changing the sets you are given. We simplified our proof in the last example by switching in the second case from working with $2^{\mathbb{N}}$ to working with $\mathcal{P}(\mathbb{Q})$, which we could do because we already knew these sets had the same size.

We end this section with a final classic fundamental theorem which you will prove in the homework.

Theorem 2.25 (Cantor's Theorem)**.** *For every set X, $|X| < |\mathcal{P}(X)|$.*

3 Well orderings

For the next two sections, we want to move away from thinking about cardinality and toward thinking about ordinal notions. The notion of cardinality concerns only the size of a set. When we think about the number 3 in terms of cardinality, we think of it denoting a set of size three, i.e. with three elements. To think about 3 *ordinally* is to think about three elements (or three people) in a queue. That is, we want to switch from thinking about sizes one, two, three and so on to thinking about positions in a queue such as first, second, third and so on. We need to set some mathematical background before getting more formally to ordinal notions.

Often in mathematics, we want to specify when two algebraic structures are essentially the same even if they are not necessary identical. The technical term is to say the structures are *isomorphic* and the definition depends on the class of structures considered. We will give a general definition for isomorphism later in the course, but for now we restrict our attention to the case of linear orders.

Definition 3.1. Let $(A, \leq_A)$ and $(B, \leq_B)$ be linear orders and let $f : A \to B$ be a function. We say f is an *isomorphism between* $(A, \leq_A)$ and $(B, \leq_B)$ if f is a bijection which preserves the orderings in the sense that for all $u, v \in A$

$$u \leq_A v \Leftrightarrow f(u) \leq_B f(v).$$

We say that $(A, \leq_A)$ and $(B, \leq_B)$ are *isomorphic*, and write $(A, \leq_A) \cong (B, \leq_B)$, if there is an isomorphism between them.

When discussing isomorphisms, or isomorphic structures, we often denote the structures by their domains and drop the associated relations. That is, we write $A \cong B$ rather than $(A, \leq_A) \cong (B, \leq_B)$ in case when the intended relations are clear.

Example 3.2. Let the linear order $(A, \leq_A)$ be given by $A = \{0, 2, 4, \ldots\}$ with $n \leq_A m$ if and only if $n \leq_{\mathbb{N}} m$. $(A, \leq_A) \cong (\mathbb{N}, \leq_{\mathbb{N}})$ by the isomorphism $f(x) = x/2$.

Example 3.3. $(\mathbb{N}, \leq_{\mathbb{N}}) \ncong (\mathbb{R}, \leq_{\mathbb{R}})$ because there is no bijection between $\mathbb{N}$ and $\mathbb{R}$. That is, $\mathbb{R}$ is too big for these linear orders to be isomorphic.

Example 3.4. It should seem intuitive that $(\mathbb{N}, \leq_{\mathbb{N}}) \ncong (\mathbb{Z}, \leq_{\mathbb{Z}})$. To show this formally, suppose for a contradiction that $f : \mathbb{N} \to \mathbb{Z}$ is an isomorphism. Because f is order preserving, we know that

$$n \leq_{\mathbb{N}} m \Leftrightarrow f(n) \leq_{\mathbb{Z}} f(m)$$

for all $n, m \in \mathbb{N}$. Let $z \in \mathbb{Z}$ be such that $z = f(0)$. To derive a contradiction, consider the element $z - 1 \in \mathbb{Z}$. What element $n \in \mathbb{N}$ satisfies $f(n) = z - 1$? Because f is onto, there must be some $n \in \mathbb{N}$ such that $f(n) = z - 1$. Because f is one-to-one and $f(0) = z$, we know that $n \neq 0$ and hence $0 <_{\mathbb{N}} n$. However, because f preserves the order on these structures, we must have $f(0) <_{\mathbb{Z}} f(n)$ which implies that $z <_{\mathbb{Z}} z - 1$ giving the desired contradiction.

These examples lead to the following observation which will emerge later in the course. If you want to show that two algebraic structures are isomorphic, you need to exhibit an isomorphism between them. (This is not entirely accurate. You might prove two structures are isomorphic by contradiction or some other nonconstructive method.) However, to show that two structures are not isomorphic, you need to show that every possible map between them fails to be an isomorphism. You might get lucky and have structures with domains of different sizes. (Why does that mean they are not isomorphic?) But, generally, things will not be that simple. Later, we will develop some tools in logic for showing that two structures are not isomorphic – and even more interestingly, see when these tools are not sufficient.

Before getting to the main algebraic objections of this section, we need to review one more piece of elementary mathematics. One of the standard methods of establishing properties of the natural number is to use induction. The method of induction on $\mathbb{N}$ works as follows. Given some property $P(x)$ which you want to show holds of every natural number, you can try to show that $P(0)$ holds (the base case) and that for every natural number n, if $P(n)$ holds, then $P(n + 1)$ holds (the induction case). If you can do these two steps, you will have established that $P(x)$ holds of every natural number.

To see why $P(n)$ holds for all $n \in \mathbb{N}$ we proceed by contradiction. Suppose $P(n)$ fails for some n. Then there is a least natural number n for which $P(n)$ fails. This number n cannot be 0 because you proved $P(0)$ holds. So, $n > 0$ and hence $n - 1$ is a natural number. But, if n is the least number for which $P(n)$ fails, then $P(n-1)$ holds. However, by the induction case, if $P(n-1)$ holds, then $P(n)$ also holds. This contradicts the assumption that $P(n)$ fails.

Our main algebraic structures for this section are well orders. The intuition is that well orders are the types of orderings on which proofs by induction work. That is, they are linear orders on which properties can be proved by induction in such a way that there cannot be "least counterexamples". We begin with the formal definition and develop properties of well orders before using them to do inductive proofs.

Definition 3.5. A *well order* is a linear order $(W, \leq_W)$ such that every nonempty subset $X \subseteq W$ has a $\leq_W$-least element. That is, if $X \subseteq W$ is nonempty, then there is an $a \in X$ such that $a \leq_W x$ for all $x \in X$.

Before turning to examples, think of a well order W as specifying a queue of people. That is, the elements of W are the people in a queue and the relation $\leq_W$ describes their relative position in the sense that $a <_W b$ means person a is closer to the front of the queue that person b. The fact that W is a well order says that if we take a nonempty collection of people out of this queue, then there is a well defined first person among the collection of people removed.

Example 3.6. Let $(L, \leq_L)$ be a finite linear order. L is a well order because any nonempty set $X \subseteq L$ has only finitely many elements, so one of these elements must be the least.

Example 3.7. The canonical example of a countable well order is $(\mathbb{N}, \leq_{\mathbb{N}})$. However, this is not the only example of a countable well order. Let ω denote something that is not a natural number, i.e. $\omega \notin \mathbb{N}$. (It doesn't matter what ω is, as long as it is not in $\mathbb{N}$.) Consider the linear order $(\mathbb{N} \cup \{\omega\}, \leq')$ where $\leq'$ is defined by

$$0 \leq' 1 \leq' 2 \leq' 3 \leq' \cdots \leq' \omega$$

That is, we place ω into the usual order on the natural numbers by making it the greatest element. We claim that this is a well order. Consider a nonempty set $X \subseteq \mathbb{N}$. If $X = \{\omega\}$, then ω is the $\leq'$-least element in X. Otherwise, $Y = X \cap \mathbb{N}$ is nonempty. The order $\leq'$ on Y is just $\leq_{\mathbb{N}}$ and hence Y has a $\leq'$-least element.

Example 3.8. Consider the ordering $\leq''$ on $\mathbb{N}$ given by

$$0 \leq'' 2 \leq'' 4 \leq'' \cdots \leq'' 1 \leq'' 3 \leq'' 5 \leq'' \cdots$$

in which we order the even in the usual way followed by the odds ordered in the usual way. To see that this is a well order, consider a nonempty set $X \subseteq \mathbb{N}$. If X contains only odd numbers, then the $\leq''$-least element of X is the $\leq_{\mathbb{N}}$-least odd number in X. On the other hand, if X contains an even number, then let $Y = X \cap \{n \in \mathbb{N} \mid n \text{ is even}\}$. Y is nonempty and the $\leq_{\mathbb{N}}$-least even number in Y is the $\leq''$-least element of X.

Example 3.9. Consider the following order on $\mathbb{N} \times \mathbb{N}$.

$$\langle n, m\rangle \leq_{\text{lex}} \langle p, q\rangle \Leftrightarrow n <_{\mathbb{N}} p \text{ or } \big(n = p \text{ and } m \leq_{\mathbb{N}} q\big)$$

This order is called the lexicographic order on $\mathbb{N} \times \mathbb{N}$. To see that this order is a well order, fix a nonempty set $X \subseteq \mathbb{N} \times \mathbb{N}$.

Let $\pi_1(X) = \{n \in \mathbb{N} \mid \exists m \in \mathbb{N}\,(\langle n, m\rangle \in X)\}$ be the projection of the elements of X onto their first coordinates. Since $\pi_1(X)$ is a nonempty subset of $\mathbb{N}$, it has a $\leq_{\mathbb{N}}$-least element a. Note that for every $\langle n, m\rangle \in X$, we know $a \leq_{\mathbb{N}} n$. Also, since $a \in \pi_1(X)$, we know that there is some $m \in \mathbb{N}$ such that $\langle a, m\rangle \in X$.

Let $Y = \{m \in \mathbb{N} \mid \langle a, m\rangle \in X\}$. Y is a nonempty subset of $\mathbb{N}$, so it has a $\leq_{\mathbb{N}}$-least element b. Note that $\langle a, b\rangle \in X$ and that for every $\langle a, m\rangle \in X$, we have $b \leq_{\mathbb{N}} m$.

We claim that $\langle a, b\rangle$ is the $\leq_{\text{lex}}$-least element of X. To see why, fix $\langle n, m\rangle \in X$ and we show that $\langle a, b\rangle \leq_{\text{lex}} \langle n, m\rangle$. We know that $a \leq_{\mathbb{N}} n$. If $a <_{\mathbb{N}} n$, then $\langle a, b\rangle \leq_{\text{lex}} \langle n, m\rangle$ because of the first components. Therefore, assume that $n = a$. In this case, we are comparing $\langle a, b\rangle$ and $\langle a, m\rangle$. However, by the previous paragraph, we know that $b \leq_{\mathbb{N}} m$ and hence $\langle a, b\rangle \leq_{\text{lex}} \langle a, m\rangle$ because of their second components.

The last three examples can be generalized into three methods for building new well orders from known well orders.

Example 3.10. Let $(W, \leq_W)$ be a well order. Fix some $a \notin W$. We can extend the well order on W by placing a as a new greatest element. That is, the linear order given by $(W \cup \{a\}, \leq')$ with

$$x \leq' y \Leftrightarrow \big(x, y \in W \text{ and } x \leq_W y\big) \text{ or } y = a$$

is a well order.

Example 3.11. Let $(A, \leq_A)$ and $(B, \leq_B)$ be well orders. We can think of "adding" these orders as follows. Intuitively, we want to put down a copy of A ordered by $\leq_A$ and then insert a copy of B ordered by $\leq_B$ so that all the elements of B come after all the elements of A. More formally, we define a well order $(A \times \{0\} \cup B \times \{1\}, \preceq)$ with

$$\langle x, y\rangle \preceq \langle u, v\rangle \preceq \Leftrightarrow y <_{\mathbb{N}} v \text{ or } \big(y = v = 0 \text{ and } x \leq_A u\big) \text{ or } \big(y = v = 1 \text{ and } x \leq_B u\big)$$

The role of the sets $\{0\}$ and $\{1\}$ is simply to separate the elements of A from the elements of B (as these sets could contain elements in common) in such a way that we can easily determine whether a element in the "sum" came from A or B. We compare the components containing a 0 or 1 first because we want to make sure an element from B is always above an element from A and because we want to compare a pair of elements from A using $\leq_A$ and a pair of elements from B using $\leq_B$.

Notice that in this example, the order in which we "added" A and B mattered. That is, we put the copy of A first and the copy of B second. We could have "added" them in the other order, but that would mean putting the copy of B first and the copy of A second. At first, it might seems like this order shouldn't matter in the end, but it does.

Suppose $A = \{a\}$ is a linear order consisting of a single element and $B = (\mathbb{N}, \leq_{\mathbb{N}})$. If we place A first and then place the elements of B greater than A, we get

$$a < 0 < 1 < 2 < 3 < \cdots$$

which is isomorphic to $(\mathbb{N}, \leq_{\mathbb{N}})$. However, if we place B first and then place the element of A greater than B, we get

$$0 < 1 < 2 < 3 < \cdots < a$$

which is not isomorphic to $(\mathbb{N}, \leq_{\mathbb{N}})$. Thus this form of "addition" is not commutative.

We can also "multiply" two linear orders. Before giving the details of this construction, let me motivate it with a simple example because this operation will also not be commutative. Consider multiplying two natural numbers such as 2×3. Think of the numbers 2 and 3 as representing ordinals (i.e. well order queues) rather than cardinalities (i.e. collections of 2 or 3 objects).

In the product 2×3, the 2 (which comes first) tells us we are considering queues each of which consists of 2 people. The 3 (which comes second) tells use that we have 3 such lines. But, remember that we want to think of the 3 in an ordinal way rather than a cardinal way, so we think of the 3 as telling us that we have a first line of 2 people, a second line of 2 people and a third line of 2 people.

Now, suppose we want to form a single line. The natural thing to do is to take the people in the first queue and place them in the new single line first (retaining their order within their line of 2 people). Then we take the people in second queue (retaining their order) and place them in the new single line after the people from the first queue. Finally, we take the people from the third queue (retaining their order) and place them at the end of the new single line.

Putting this altogether, we think of 2×3 as combining 3 queues of 2 people into a single line by place each of the 3 queues into the new single line in order (of the queues as given by the 3) and retaining the given order within each queue. Try to keep this picture in mind for the next example.

Example 3.12. Let $(A, \leq_A)$ and $(B, \leq_B)$ be well orders. We can think of "multiplying" these orders as follows. Intuitively, we start with a bunch of queues isomorphic to A. The collection of these queues is ordered as described by B. We want to combine the A-queues into a new single line. The order B on the collection of the A-queues tells us which order to place the A-queues in the new single line.

Taking one step towards formality, think of each element $b \in B$ as denoting the b-th A-queue. When we assemble the A-queues into a single line, we want to replace b by a copy of A (i.e. by the b-th A-queue). That is, we want to take the well order B and replace each element of B by a copy of A. If $a \in A$ and $b \in B$, we represent the person in the b-th queue and in the a-th position within that queue by $\langle b, a \rangle$. In other words, the "address" for each person in the new line is given by specifying which queue they came from (i.e. the b-th queue) followed by their position within that queue (the a-th spot in that queue).

Formally, we define the linear order $(B \times A, \leq_{\text{lex}})$ with

$$\langle b, a \rangle \leq_{\text{lex}} \langle d, c \rangle \Leftrightarrow b <_B d \text{ or } \big(b = d \text{ and } a \leq_A c\big)$$

This order is called the *lexicographic order* on $B \times A$. Notice that we compare two people in the new single line by first comparing which A-queue they originally came from and then, if they can from the same A-queue, comparing their positions within that A-queue.

Notice that commutativity fails for this notion of multiplication. Suppose $A = \{a, b\}$ with $a <_A b$ is a well order with two elements and $B = (\mathbb{N}, \leq_{\mathbb{N}})$. If we multiply A by B, then we have a collection of 2 person queues ordered by $\mathbb{N}$. Let $A_n = \{a_n, b_n\}$ be the n-th queue in this $\mathbb{N}$-order of queues. When we combines the 2 person queues into a single line, we get

$$a_0 < b_0 < a_1 < b_1 < a_2 < b_2 < \cdots$$

which is isomorphic to $(\mathbb{N}, \leq_{\mathbb{N}})$. If we multiply B by A, then we have two queues each looking like $\mathbb{N}$. Let $\{0_a, 1_a, \ldots\}$ be the a-th $\mathbb{N}$-queue and let $\{0_b, 1_b, \ldots\}$ be the b-th $\mathbb{N}$-queue. When we combine the $\mathbb{N}$-queues into a single line, we get

$$0_a < 1_a < 2_a < \cdots < 0_b < 1_b < 2_b < \cdots$$

which is not isomorphic to $(\mathbb{N}, \leq_{\mathbb{N}})$. It is not a coincidence that it is isomorphic to the well order we get by "adding" to copies of $(\mathbb{N}, \leq_{\mathbb{N}})$. We will see this connection in detail later.

These examples of adding and multiplying will return when we consider ordinal numbers. Next, we give a simple characterization of when a linear order is a well order.

Lemma 3.13. *A linear order $(L, \leq_L)$ is a well order if and only if L has no infinite descending sequences.*

Proof. We can state this lemma equivalently as a linear order is not a well order if and only if it has an infinite descending sequence. It is easier to prove in this form.

First, suppose that L is not a well order and we construct an infinite descending sequence. Since L is not a well order, there is a nonempty set $X \subseteq L$ such that X has no $\leq_L$-least element. Define the infinite descending sequence $s : \mathbb{N} \to L$ as follows. Set $s(0) = x_0$ where x_0 is any element of X. Since X has no $\leq_L$-least element, we know there is an element $x_1 \in X$ such that $x_1 <_L x_0$. Fix such an x_1 and set $s(1) = x_1$. Now, repeat this process. That is, assume that $s(n) = x_n \in X$ has been defined. There must be an element $x_{n+1} \in X$ such that $x_{n+1} <_L x_n$. Fix such that element and set $s(n+1) = x_{n+1}$. By definition, s is an infinite descending sequence in X and hence in L.

Second, suppose that $s : \mathbb{N} \to L$ is an infinite descending sequence and we show that L is not a well order. Let $X = \text{range}(s)$. Note that X is nonempty. To see that X has no $\leq_L$-least element, fix $x \in X$. By the definition of X, there is an $n \in \mathbb{N}$ such that $s(n) = x$. Since $s(n+1) \in X$ and $s(n+1) <_L s(n) = x$, the element x is not $\leq_L$-least in X. Therefore, X witnesses that L is not a well order. □

For the remainder of this section, we will consider some special properties of well orders which are not generally shared by linear orders. The first special property of well orders we will consider is the representation of initial segments in linear orders.

Definition 3.14. Let L be a linear order. A subset $I \subseteq L$ is an *initial segment of* L if it satisfies

$$\forall x, y \left((x \in I \wedge y \leq_L x) \to y \in I \right)$$

In other words, I is an initial segment of L if it is closed downwards under $\leq_L$. An initial segment I is called *proper* of $I \neq L$.

Example 3.15. Let $(L, \leq_L)$ be a linear order. $L \subseteq L$ is an initial segment of itself and $\emptyset \subseteq L$ is an initial segment. Unless L is empty, $\emptyset$ is a proper initial segment of L.

Definition 3.16. Let L be a linear order and $a \in L$. The *initial segment generated by* a is

$$I(a) = \{x \in L \mid x <_L a\}.$$

Note the strict inequality in this definition. (You should check that $I(a)$ is an initial segment.)

Example 3.17. Consider the linear order $(\mathbb{Q}, \leq_{\mathbb{Q}})$ and the initial segment

$$I = \{q \in \mathbb{Q} \mid q \leq_{\mathbb{Q}} 0\}.$$

Notice that I is not generated by any element of $\mathbb{Q}$. That is, to try to find a q such that $I = I_q$, you would have to have $0 <_{\mathbb{Q}} q$ because $0 \in I$. However, if $0 <_{\mathbb{Q}} q$, then $0 <_{\mathbb{Q}} q/2 <_{\mathbb{Q}} q$ and hence $q/2 \in I(q)$ but $q/2 \notin I$. Therefore, $I \neq I_q$.

For a more subtle example, consider the initial segment $I' = \{q \in \mathbb{Q} \mid q <_{\mathbb{R}} \sqrt{2}\}$. Because $\sqrt{2} \notin \mathbb{Q}$, this initial segment is also not generated by any element of $\mathbb{Q}$.

This example shows that for general linear orders, we cannot expect initial segments to be generated by individual elements. Even worse, a linear like $(\mathbb{Q}, \leq_{\mathbb{Q}})$ has uncountably many district initial segments (given by the Dedekind cuts) but has only countably many initial segments generated by its elements! The next lemma shows that this behavior cannot occur in a well order.

Lemma 3.18. *Let W be a well order. $I \subseteq W$ is a proper initial segment if and only if there is an element $a \in W$ such that $I = I(a)$.*

Proof. Since I is proper, $W \setminus I \neq \emptyset$. Therefore, because W is a well order, $W \setminus I$ has a $\leq_W$-least element a. We claim that $I = I(a)$. To prove this claim, we show that $I \subseteq I(a)$ and $I(a) \subseteq I$.

To see that $I \subseteq I(a)$, assume for a contradiction that $I \not\subseteq I(a)$. This means that there is an element $b \in I$ such that $b \notin I(a)$ and hence $a \leq_W b$. However, I is an initial segment, so $a \leq_W b$ and $b \in W$ implies that $a \in I$. This contradicts the fact that $a \in W \setminus I$.

To see that $I(a) \subseteq I$, assume for a contradiction that $I(a) \not\subseteq I$. This means that there is an element $c \in I(a)$ such that $c \notin I$. Since $c \in I(a)$, we have $c <_W a$. Since $c \notin I$, we have $c \in W \setminus I$ and thus, because a is the $\leq_W$-least element of $W \setminus I$, $a \leq_W c$. Having shown $c <_W a$ and $a \leq_W c$, we have arrived at the desired contradiction. □

The second property of well orders we consider corresponds to strong induction on $\mathbb{N}$. We will use this property repeatedly in the rest of the section.

Theorem 3.19 (Transfinite Induction)**.** *Let W be a well order and let $B \subseteq W$. If for every $x \in W$, we have $I(x) \subseteq B \Rightarrow x \in B$, then $B = W$.*

Proof. For a contradiction, assume that $B \subsetneq W$ but

$$\forall x \in W \big(I(x) \subseteq B \to x \in B\big).$$

Since $B \subsetneq W$, we know that $W \setminus B \neq \emptyset$ and thus $W \setminus B$ has a $\leq_W$ least element a. However, if a is the $\leq_W$-least element of $W \setminus B$, then $I(a) \subseteq B$. Therefore, by the offset equation, $a \in B$ for the desired contradiction. □

Definition 3.20. Let A and B be linear orders and let $f : A \to B$ be a function between them. We say

- f is *increasing* if $x \leq_A y$ implies $f(x) \leq_B f(y)$ for all $x, y \in A$, and
- f is *strictly increasing* if $x <_A y$ implies $f(x) <_B f(y)$ for all $x, y \in A$.

Example 3.21. Let $(A, \leq_A)$ and $(B, \leq_B)$ be linear orders and let $f : A \to B$ be an isomorphism between them. By the definition of isomorphism, for all $x, y \in A$, we have $x \leq_A y$ if and only if $f(x) \leq_B f(y)$. Therefore, f is increasing.

In fact, if f is an isomorphism, we have the stronger property that $x <_A y$ if and only if $f(x) <_B f(y)$ because f is one-to-one. Therefore, f is also strictly increasing.

Example 3.22. A function $f : A \to B$ between linear orders does not have to be an isomorphism to be increasing or strictly increasing. For example, any constant function is increasing. To give another example, consider $f : \mathbb{N} \to \mathbb{N}$ given by $f(x) = 2x$. Then f is increasing but is not an isomorphism because it is not onto.

Exercise 3.23. Show that if $f : A \to B$ is strictly increasing, then f is one-to-one. We will use this property several times below.

Theorem 3.24. *Let W be a well order. If $f : W \to W$ is strictly increasing, then $x \leq_W f(x)$ for all $x \in W$.*

Before proving this theorem, notice that it is not true about linear orders in general. Let $f : \mathbb{Z} \to \mathbb{Z}$ be defined by $f(x) = x - 1$. Then f is strictly increasing because $x <_{\mathbb{Z}} y$ implies $f(x) = x - 1 <_{\mathbb{Z}} y - 1 = f(y)$. However, we have $f(x) <_{\mathbb{Z}} x$ for all $x \in \mathbb{Z}$.

We will give two proofs of Theorem 3.24 to illustrate two different ways in which proofs by transfinite induction are commonly given.

First proof of Theorem 3.24. For this proof, we apply transfinite induction directly. Let

$$B = \{z \in W \mid z \leq_W f(z)\}$$

We want to show that $B = W$. By transfinite induction, it suffices to show that for all $x \in W$, if $I(x) \subseteq B$, then $x \in B$. Fix $x \in W$ and assume that $I(x) \subseteq B$. We will show that $x \in B$.

We claim that

$$y \in I(x) \text{ implies that } y <_W f(x).$$

To prove this claim, fix an element $y \in I(x)$. First, notice that since $y \in I(x)$ and $I(x) \subseteq B$, we have $y \in B$ which means that $y \leq_W f(y)$. Second, notice that $y \in I(x)$ implies $y <_W x$ (by definition), which in turn implies that $f(y) <_W f(x)$ because f is strictly increasing. Putting these inequalities together, we have that $y \leq_W f(y) <_W f(x)$. This completes the proof of the claim.

Because $f(x) \not<_W f(x)$, this claim implies that $f(x) \notin I(x)$. (Substitute $f(x)$ in for y in the claim.) However, by the definition of the initial segment $I(x)$, $f(x) \notin I(x)$ means that $x \leq_W f(x)$. □

Second proof of Theorem 3.24. Probably the more common way in which proofs by transfinite induction proceed is by contradiction. Assume for a contradiction that

$$C = \{z \in W \mid f(z) <_W z\} \neq \emptyset$$

Since W is a well order, we can fix the $\leq_W$-least element $c \in C$. To derive our contradiction, we need two observations.

First, since $c \in C$, we have $f(c) <_W c$. Applying f to both sides of this inequality, we have $f(f(c)) <_W f(c)$ since f is strictly increasing.

Second, since c is the $\leq_W$-least element of C and $f(c) <_W c$, we know that $f(c) \notin C$. Therefore, by the definition of C, we have $f(f(c)) \not<_W f(c)$. Thus, we have the desired contradiction. □

Corollary 3.25. *If W is a well order and $x \in W$, then $W \ncong I(x)$.*

Proof. Suppose for a contradiction that there is an $x \in W$ and a map $f : W \to I(x)$ that is an isomorphism. On one hand, we know that $f(x) \in I(x)$ since f maps into $I(x)$ and hence $f(x) <_W x$. On the other hand, f is an isomorphism between linear orders, so it is strictly increasing by Example 3.21. Therefore, by Theorem 3.24, $x \leq_W f(x)$ giving the desired contradiction. □

Corollary 3.26. *If A is a well order and $x \neq y \in A$, then $I(x) \ncong I(y)$.*

Proof. Suppose for a contradiction that $x \neq y \in A$ but $I(x) \cong I(y)$. Without loss of generality, assume that $x <_A y$. Let $W = I(y)$ be the well order given by the initial segment determined by y. Since $x <_A y$, we have $x \in W$ and we can consider $I(x)$ as an initial segment of W. Therefore, we have $W \cong I(x)$ where $x \in W$. This directly contradicts Corollary 3.25. □

The third general property we want to consider about well orders is that isomorophisms between them are unique. Again, this property fails badly for linear orders in general. Consider $\mathbb{Z}$. Every function $f : \mathbb{Z} \to \mathbb{Z}$ of the form $f(x) = x + n$ for $n \in \mathbb{Z}$ is an isomorphism.

Theorem 3.27. *Let A and B be isomorphic well orders. There is a unique isomorphism from A to B.*

Proof. Suppose for a contradiction that there are two different isomorphisms $f : A \to B$ and $g : A \to B$. Saying that these two isomorphisms are different means that there is some $x \in A$ such that $f(x) \neq g(x)$. In other words,

$$C = \{x \in A \mid f(x) \neq g(x)\} \neq \emptyset.$$

Since A is a well order, we can fix the $\leq_A$-least element $c \in C$. Because $c \in C$, we know $f(c) \neq g(c)$ (as elements of the linear order B), so without loss of generality, we assume that $f(c) <_B g(c)$. (That is, either $f(c) <_B g(c)$ or $g(c) <_B f(c)$, and we assume that the functions are named so that $f(c)$ is the smaller element in B.)

Let $b = f(c)$. We have two important pieces of information. First, $b <_B g(c)$. Second, for all $x <_A c$, $f(x) = g(x)$ because c is the $\leq_A$-least element of C. Now, we ask the crucial question: For which $a \in A$ does $g(a) = b$? There must be some such a because g is an isomorphism. There are three possibilities (either $a <_A c$ or $a = c$ or $c <_A a$) and we show that none of them is actually possible.

- $a <_A c$: In this case, $g(a) = f(a)$ by the second piece of information above. Because f is an isomorphism and hence is strictly increasing, $a <_A c$ implies $f(a) <_B f(c)$. Therefore, we have $g(a) = f(a) <_B f(c) = b$, so $g(a) <_B b$ and hence $g(a) \neq b$.
- $a = c$: In this case, $g(a) = g(c) >_B b$ (by the first piece of information) and hence $g(a) \neq b$.
- $c <_A a$: Because g is an isomorphism and hence is strictly increasing, $c <_A a$ implies $g(c) <_B g(a)$. Therefore, $b <_B g(c) <_B g(a)$, so $b <_B g(a)$ and hence $g(a) \neq b$.

Having analyzed these three possibilities, we conclude that there is no element $a \in A$ such that $g(a) = b$ and hence we have contradicted the fact that g is an isomorphism. □

For the final theorem of this section, it will be important to keep track of initial segments in two different well orders. To help keep everything straight, if $(A, \leq_A)$ is a linear order and $a \in A$, we use $I_A(a)$ to denote the initial segment determined by a in A. The subscript A is implied by the context, but it makes things somewhat more clear to denote it explicitly. The following technical fact will be useful in our final theorem of this sections.

Lemma 3.28. *Let A and B be isomorphic well orders and let $f : A \to B$ be the isomorphism between them. Fix $a \in A$ and let $b = f(a)$. Then $I_A(a) \cong I_B(b)$ by the function $f \restriction I_A(a)$.*

Proof. Recall that $f \restriction I_A(a)$ is the restriction of f to $I_A(a)$. That is, the domain of $f \restriction I_A(a)$ is $I_A(a)$ and for each $c \in I_A(a)$, $(f \restriction I_A(a))(c) = f(c)$. To prove this lemma, we verify the required properties of $f \restriction I_A(a)$.

First, we claim that $f \restriction I_A(a)$ maps $I_A(a)$ into $I_B(b)$. Fix $c \in I_A(a)$ and we show that $(f \restriction I_A(a))(c) \in I_B(b)$. Since $c \in I_A(a)$, we have $c <_A a$. Because f is an isomorphism, this inequality implies that $f(c) <_B f(a) = b$ and hence $f(c) \in I_B(b)$. But, $(f \restriction I_A(a))(c) = f(c)$, so $(f \restriction I_A(a))(c) \in I_B(b)$.

Second, we claim that $f \restriction I_A(a)$ is one-to-one and order preserving. Both of these facts follow immediately because $(f \restriction I_A(a))(c) = f(c)$ and f is one-to-one and order preserving.

Third, we claim that $f \restriction I_A(a)$ is onto $I_B(b)$. Fix $d \in I_B(b)$. Because f is an isomorphism, and hence is a bijection, there is an element $c \in A$ such that $f(c) = d$. Because f is order preserving and $d <_B b$, we must have $c <_A a$ and hence $c \in I_A(a)$. But, then $(f \restriction I_A(a))(c) = f(c) = d$ as required.

We have shown that $f \restriction I_A(a)$ is an order preserving bijection between $I_A(a)$ and $I_B(b)$. Hence, by definition, it is an isomorphism between the well orders $I_A(a)$ and $I_B(b)$. □

Finally, we reach the most important property of well orders. The final theorem of this section shows that we can always compare the lengths of well orders. That is, given two well orders, either they have the same length (i.e. are isomorphic) or one is strictly short than the other (i.e. isomorphic to an initial segment of the other).

Theorem 3.29. *Let A and B be well orders. Exactly one of the following holds.*

(1) $A \cong B$

(2) $A \cong I(b)$ *for some* $b \in B$

(3) $B \cong I(a)$ *for some* $a \in A$

Proof. For this proof, it is important to keep track of initial segments of A and initial segments of B. To emphasize which well order we are working in, if $a \in A$, I will use $I_A(a)$ to indicate that the initial segment we are looking at is in A. Similarly, if $b \in B$, I will use $I_B(b)$ to indicate that the initial segment is in B.

We begin by showing that at least one of these conditions must hold. Define the following set Z.

$$Z = \{x \in A \mid \exists y \in B \left(I_A(x) \cong I_B(y)\right)\}$$

Our proof will follow from a series of claims.

Our first claim is that if $x \in Z$, then there is a unique $y \in B$ such that $I_A(x) \cong I_B(y)$. Suppose for a contradiction that there are $y_1 \neq y_2 \in B$ such that $I_A(x) \cong I_B(y_1)$ and $I_A(x) \cong I_B(y_2)$. Then $I_B(y_1) \cong I_B(y_2)$ which contradicts Corollary 3.26. This proves the first claim.

By this first claim, we can define a function $f : Z \to B$ by

$$f(x) = y \Leftrightarrow I_A(x) \cong I_B(y).$$

Our second claim is that Z is an initial segment of A. To see why, fix $z \in Z$ and $x \in A$ with $x <_A z$. We need to show that $x \in Z$. That is, we need to find an element $y \in B$ such that $I_A(x) \cong I_B(y)$. Since $z \in Z$, we have $I_A(z) \cong I_B(f(z))$ and can fix an isomorphism $h : I_A(z) \to I_B(f(z))$. Since $x <_A z$, the element x is in the domain of h. Let $y = h(x)$. By Lemma 3.28, the restricted function $h \restriction I_A(x)$ is an isomorphism from $I_A(x)$ to $I_B(y)$. completing the proof of the second claim.

In particular, notice that in the context of the second claim, $f(x) = y$ because $I_A(x) \cong I_B(y)$. Since $y = h(x)$, we have $f(x) = h(x)$. That is, if $z \in Z$, $x <_A z$ and $h : I_A(z) \to I_B(f(z))$ is the isomorphism between $I_A(z)$ and $I_B(f(z))$, then $f(x) = h(x)$.

Our third claim is that f is strictly increasing. To see why, fix $x, z \in Z$ with $x <_A z$. Since $z \in Z$, we can fix an isomorphism $h : I_A(z) \to I_B(f(z))$. By the comment in the previous paragraph, $f(x) = h(x) \in I_B(f(z))$. This means that $f(x) <_B f(z)$ as required.

Let $Y \subseteq B$ be the image of Z under f. Notice that $f : Z \to Y$ is a function (by the first claim) which is onto (by the definition of Y) and which is one-to-one (because f is strictly increasing). Therefore, f is a bijection.

Our fourth claim is that Y is an initial segment of B. To see why, fix $b \in Y$ and $b' <_B b$. We need to show that $b' \in Y$. Fix $a \in Z$ such that $f(a) = b$ and fix the isomorphism $h : I_A(a) \to I_B(b)$. Since h maps $I_A(a)$ onto $I_B(b)$ and since $b' \in I_B(b)$, there is an $a' \in I_A(a)$ such that $h(a') = b'$. By Lemma 3.28, $h \restriction I_A(a')$ is an isomorphism from $I_A(a')$ to $I_B(b')$. Therefore, $f(a') = b'$ and $b' \in Y$.

We have arrived at a bijection $f : Z \to Y$ between the initial segment Z of A and the initial segment Y of B. Because f is strictly increasing, it is order preserving. Therefore, f is an isomorphism between the initial segment $(Z, \leq_A)$ of A and the initial segment $(Y, \leq_B)$ of B. We split into four cases to finish the proof.

- Suppose $Z = A$ and $Y = B$. Then $f : A \to B$ is an isomorphism between A and B.
- Suppose Z is a proper initial segment of A and $Y = B$. Then $Z = I_A(a)$ for some $a \in A$ and $f : I_A(a) \to B$ is an isomorphism.
- Suppose $Z = A$ and Y is a proper initial segment of B. Then $Y = I_B(b)$ for some $b \in B$ and $f : A \to I_B(b)$ is an isomorphism.
- Suppose Z is a proper initial segment of A and Y is a proper initial segment of B. In this case, $Z = I_A(a)$ for some $a \in A$, $Y = I_B(b)$ for some $b \in B$ and $f : I_A(a) \to I_B(b)$ is an isomorphism. But, then $I_A(a) \cong I_B(b)$ and hence $a \in Z$, which means $a \in I_A(a)$ which is a contradiction. Therefore, this final case cannot occur.

Since the last case cannot occur, we must be in one of the first three cases which finished the proof that at least one of the three conditions in the statement of the Theorem must hold.

It remains to show that we cannot have two of these conditions hold. If (1) and (2) hold, then by composing the isomorphisms, we would have $B \cong I_B(b)$ for some $b \in B$ which contradicts Corollary 3.25. Similar, we cannot have (1) and (3) hold. Finally, suppose that (2) and (3) hold. Fix the isomorphisms $f : A \to I_B(b)$ and $g : B \to I_A(a)$. Let $a' = g(b)$. By Lemma 3.28, $g \restriction I_B(b)$ is an isomorphism from $I_B(b)$ to $I_A(a')$. Composing f and $g \restriction I_B(b)$, we get that $A \cong I_A(a')$ which again contradicts Corollary 3.25. □

4 Ordinal numbers

In this section, we outline the theory of the ordinal numbers. The basic idea is to pick a unique representative for each well order type in a canonical manner. Later in the course, we will say something about the axioms of Zermelo-Frankel set theory with choice (ZFC) which is the natural formal theory to develop the theory of ordinal numbers.

There are two aspects of ZFC which are useful to keep in mind for this section in terms of helping with the intuition. First, you should think of everything as a set. That is, the elements of sets are other sets. Second, in ZFC, no set is allowed to be a member of itself. That is, we never have $x \in x$. In ZFC, this property is guaranteed by the Axiom of Foundation. It is not necessary have the Axiom of Foundation to develop ordinal numbers, but it is helpful when you are first thinking about them. So, although we will not use it formally, it doesn't hurt to read this material under the working hypothesis that any set x, $x \notin x$.

Definition 4.1. A set A is *transitive* if for every $x \in A$, we have $x \subseteq A$. That is,

$$\forall x \, \forall w \, \big((w \in x \wedge x \in A) \rightarrow w \in A \big).$$

Example 4.2. $\emptyset$ is a transitive set. The set $\{\emptyset\}$ is also transitive because it has one element, namely $\emptyset$, and $\emptyset \subseteq \{\emptyset\}$.

The set $A = \{\emptyset, \{\emptyset\}\}$ is also transitive. To check this, we consider each of the elements of A in turn. First, $\emptyset \in A$ and $\emptyset \subseteq A$. Second, $\{\emptyset\} \in A$ and $\{\emptyset\} \subseteq A$ because $\emptyset \in A$.

However, the set $B = \{\{\emptyset\}\}$ is not transitive because $\{\emptyset\} \in B$ but $\{\emptyset\} \not\subseteq B$ because $\emptyset \notin B$.

Definition 4.3. A set α is an *ordinal* (or an *ordinal number*) if α is transitive and the membership relation $\in$ defines a strict well order on α. That is, it defines a strict linear order which is a well order.

We tend the drop the adjective "strict" when talking about the $\in$-relation on an ordinal and just refer to the ordinal as being well ordered by the $\in$-relation. However, when checking that a given set is an ordinal, it is important to keep in mind that the axioms we need to check are those of a strict linear order. We begin with a collection of examples that connect ordinals with the natural numbers.

Example 4.4. The simplest ordinal is $\emptyset$ which corresponds to a well order with no elements. For this reason, we denote this ordinal by 0.

Example 4.5. The next simplest ordinal is $\{\emptyset\}$. We have already seen that $\{\emptyset\}$ is a transitive set. It remains to show that the $\in$-relation defines a well order on $\{\emptyset\}$.

The set $\{\emptyset\}$ has a single element, namely $\emptyset$. Since $\emptyset \notin \emptyset$ (because $\emptyset$ has no elements), the $\in$-relation on $\{\emptyset\}$ is irreflexive. It also satisfies transitivity and trichotomy trivially, so defines a strict linear order on $\{\emptyset\}$. Because this linear order has only one element, it is a well order. Therefore, $\{\emptyset\}$ is an ordinal. Because this ordinal corresponds to a well order with one element, so we denote it by 1. There are two important points to make about this ordinal. First, because $0 = \emptyset$, we have $1 = \{0\}$ and in particular $0 \in 1$.

Second, consider whether we could have another ordinal with a single element. Suppose $A = \{a\}$ is an ordinal and $A \neq 1$. Then, in particular, $a \neq \emptyset$. If A is an ordinal, then A is transitive so $a \in A$ implies that $a \subseteq A$. Therefore, $a \subseteq \{a\}$. This means that if $b \in a$, then $b \in \{a\}$ and hence $b = a$. Therefore, a must contain an element because $a \neq \emptyset$ and the only element it can contain is a. Therefore, $a \in a$ which we have forbidden because the $\in$-relation on an ordinal satisfies the axioms of a strict linear order and hence is irrelexive. Hence 1 is the only ordinal with one element.

Our next example will set the pattern for examples corresponding to each of the natural numbers.

Example 4.6. Let $2 = \{0, 1\}$. That is, $2 = \{\emptyset, \{\emptyset\}\}$. We have already checked that 2 is transitive. The $\in$-relation on 2 holds only for $\emptyset \in \{\emptyset\}$ (i.e. $0 \in 1$). Therefore, it defines a strict linear order with two elements which is necessarily a well order because it is finite.

Notice that $0 \in 2$ and $1 \in 2$. I will leave it to you to check, but 2 is the only possible ordinal with exactly two elements.

To get used to this notation and see what exactly is going on, it is worth writing down the exact set definition of $3 = \{0, 1, 2\}$ and checking that it is transitive and that the $\in$-relation defines a well order on 3.

Example 4.7. We can now continue to define ordinals corresponding to each of the natural numbers. Assume $n = \{0, 1, \ldots, n-1\}$ has been defined with the $\in$-relation holding between $i, j \in n$ if and only if $i <_{\mathbb{N}} j$. Define $n + 1 = \{0, 1, \ldots, n-1, n\}$. Because $i \in n$ for all $i <_{\mathbb{N}} n$ and $n \notin n$, the $\in$-relation satisfies the axioms for a strict linear order on $n + 1$. Again, this linear order is a well order because it is finite, in fact it contains exactly $n+1$ many elements.

Another way to describe the general pattern of the last example is to define $n+1 = n \cup \{n\}$. Later we will show this pattern generalizes to all ordinals by defining a successor operation $S(x)$ on the ordinals by $S(\alpha) = \alpha \cup \{\alpha\}$ and verifying carefully that if α is an ordinal then $S(\alpha)$ is also an ordinal. The process of going from α to $S(\alpha)$ exactly corresponds to the process of adding a new greatest element to a well order.

Our examples so far have only generated finite ordinals. There are also infinite ordinals.

Example 4.8. Let $\omega = \{0, 1, 2, \ldots\}$ where each n is viewed in its ordinal form (i.e. as a set containing the strictly smaller ordinals). Based on the previous pattern, we have that for all $n, m \in \omega$, $n \in m$ if and only if $n <_{\mathbb{N}} m$. Therefore, the $\in$-relation defines a strict linear order on ω such that $(\omega, \in) \cong (\mathbb{N}, <_{\mathbb{N}})$. That is, ω is an ordinal notation for the well order $\mathbb{N}$ (with its usual ordering).

Example 4.9. Let $S(\omega) = \omega \cup \{\omega\} = \{0, 1, 2, \ldots\} \cup \{\omega\}$. If $x, y \in S(\omega)$, then $x \in y$ if and only if $x, y \in \omega$ and $x <_{\mathbb{N}} y$ or $y = \omega$. That is, we have

$$0 \in 1 \in 2 \in \cdots \in \omega$$

This definition should look familiar. It is exactly the process we used to add a new greatest element to the end of a well order. Therefore, $S(\omega)$ is an ordinal corresponding to taking $\mathbb{N}$ and adding a new greatest element.

Having given these motivating examples, we turn to proving basic properties about ordinals. Our eventual goal is to show that the ordinals give unique representatives of each well order isomorphism type. That is, for each (strict) well order $(W, <_W)$, there is a unique ordinal α such that $(W, <_W) \cong (\alpha, \in)$.

Before proving these facts about ordinal numbers, we should unpack the definition of an ordinal to see exactly what we need to verify. To be an ordinal, the pair $(\alpha, \in)$ should satisfy

the axioms of a strict linear order and this order should be a well order. (We will stop writing the binary relation $\in$ defining the strict linear order on α. When dealing with ordinals, the ordering is always given by $\in$.) That is, α should satisfy the following properties.

- Irreflexivity: For all $x \in \alpha$, $x \notin x$.
- Trichotomy: For all $x, y \in \alpha$, either $x = y$ or $x \in y$ or $y \in x$.
- Transitivity: For all $x, y, z \in \alpha$, if $z \in y$ and $y \in x$, then $z \in x$.
- Well order: If $B \subseteq \alpha$ is nonempty, then there is a $b \in B$ such that for all $x \in B$, $x \notin b$. In other words, there is a $b \in B$ such that $b \cap B = \emptyset$.

Lemma 4.10. *Let α be an ordinal. If $\beta \in \alpha$, then β is an ordinal and $\beta = I_\alpha(\beta)$.*

Proof. First, we show that β is a transitive set. Fix arbitrary $x \in y$ and $y \in \beta$. We need to show that $x \in \beta$. Since α is an ordinal and $\beta \in \alpha$, we know that $\beta \subseteq \alpha$. Therefore, $y \in \beta$ implies that $y \in \alpha$. Now, we can use the fact that α is an ordinal again to conclude that $y \subseteq \alpha$ and hence $x \in \alpha$.

At this point, we know that $x, y, \beta \in \alpha$ with $x \in y$ and $y \in \beta$. But, because α is an ordinal, the $\in$-relation defines a strict linear order on α. Therefore, $x \in y$ and $y \in \beta$ implies that $x \in \beta$ as required. This completes the proof that β is a transitive set.

Second, we show that the $\in$-relation defines a well order on β. Because $\beta \subseteq \alpha$ and the membership relation defines a well order on α, it automatically defines a well order on β. That is, if $x, y, z \in \beta$, then $x, y, z \in \alpha$ and hence satisfy irreflexivity, trichotomy and transitivity as elements of α. Similarly, if $B \subseteq \beta$ is nonempty, then B is also a nonempty subset of α and hence has a minimal element b under the $\in$-relation as required.

Third, we show that $\beta = I_\alpha(\beta)$. By definition,

$$I_\alpha(\beta) = \{x \in \alpha \mid x \in \beta\}$$

Therefore, $I_\alpha(\beta) \subseteq \beta$ by definition. To see $\beta \subseteq I_\alpha(\beta)$, fix $x \in \beta$. Since $\beta \subseteq \alpha$, we have $x \in \alpha$ and hence $x \in I_\alpha(\beta)$ as required. □

Lemma 4.11. *For any ordinal α, $\alpha \notin \alpha$.*

Proof. As mentioned at the beginning of this section, if we are working in ZFC and have the Axiom of Foundation, then this property follows immediately because $x \notin x$ for all sets x. However, we can get by without appealing to the Axiom of Foundation. Suppose for a contradiction that α is an ordinal and $\alpha \in \alpha$. Then $\alpha = \beta$ for some $\beta \in \alpha$. However, if $\alpha = \beta$ and $\beta \in \alpha$, then $\beta \in \beta$ by substitution. This fact contradicts the irreflexivity of the $\in$-relation on α. □

Lemma 4.12. *If α and β are ordinals such that $\alpha \cong \beta$ as well orderings, then $\alpha = \beta$.*

Proof. Assume for a contradiction that $\alpha \cong \beta$ but $\alpha \neq \beta$. Since $\alpha \neq \beta$, one of these sets must contain an element not contained in the other. Without loss of generality, assume that there is an $x \in \alpha$ such that $x \notin \beta$. By assumption, the set

$$X = \{x \in \alpha \mid x \notin \beta\} \subseteq \alpha$$

is not empty and hence has an $\in$-least element a. That is, $a \in \alpha$ and $a \notin \beta$. Furthermore, since $a \in \alpha$ and α is an ordinal, we know that $a \subseteq \alpha$. Therefore, for every $v \in a$, we have $v \in \alpha$. Because a is the $\in$-least element in X, we must have that each $v \in a$ is also in β. Therefore, $a \subseteq \beta$ but $a \notin \beta$.

Next, notice that $a = I_\alpha(a)$ is an initial segment of α by Lemma 4.10. Also, by Lemma 4.10, a is an ordinal and hence is a transitive set. We claim that $a \subseteq \beta$ is also an initial segment of β under the $\in$-relation. To prove this claim, we need to show that if $v \in a$ and $u \in v$, then $u \in a$. However, this property immediately holds because a is transitive. Therefore, $a \subseteq \beta$ is an initial segment of β.

Let $f : \alpha \to \beta$ be the unique isomorphism between the well orders α and β and recall how this isomorphism is built in Theorem 3.29. Because the isomorphism is unique, for any $x \in \alpha$ and $y \in \beta$, we must have $f(x) = y$ if and only if $I_\alpha(x) \cong I_\beta(y)$. Consider any element $v \in a$. We know $v \in \alpha$ and $v \in \beta$. Therefore, by Lemma 4.10, $v = I_\alpha(v) = I_\beta(v)$ and hence $I_\alpha(v) \cong I_\beta(v)$ by the identity map. Therefore, $f(v) = v$ for all $v \in a$.

However, we know that $a \notin \beta$. Therefore, $f(a) \in \beta$ but $f(a) \notin a$ because $f(v) = v$ for all $v \in a$. This means that $a \subseteq \beta$ is a proper initial segment of β because it does not contain $f(a)$. By Lemma 3.18, there is an element $b \in \beta$ such that $a = I_\beta(b)$. But, by Lemma 4.10, $b = I_\beta(b)$ and hence $a = b$. Thus, $a \in \beta$ which contradicts the fact that $a \in X$. □

Lemma 4.13. *If α and β are ordinals, then either $\alpha \in \beta$ or $\alpha = \beta$ or $\beta \in \alpha$.*

Proof. Fix ordinals α and β. By Theorem 3.29, we know that one of the following three cases must hold.

- Suppose $\alpha \cong \beta$. By Lemma 4.12, $\alpha = \beta$.
- Suppose $\alpha \cong I_\beta(b)$ for some $b \in \beta$. By Lemma 4.10, b is an ordinal and $b = I_\beta(b)$. Therefore, $\alpha \cong b$ and by Lemma 4.12, $\alpha = b$. But, then $\alpha \in \beta$ because $b \in \beta$.
- Suppose $I_\alpha(a) \cong \beta$ for some $a \in \alpha$. By the same reasoning as the previous case, $\beta \in \alpha$.

□

Lemma 4.14. *If α, β and γ are ordinals with $\alpha \in \beta$ and $\beta \in \gamma$, then $\alpha \in \gamma$.*

Proof. Since γ is an ordinal, and hence is transitive, $\beta \in \gamma$ implies $\beta \subseteq \gamma$. Therefore, $\alpha \in \beta$ implies $\alpha \in \gamma$. □

Lemma 4.15. *If C is a nonempty set (or class) of ordinals, then there is an $\alpha \in C$ such that for all $\beta \in C$, either $\beta = \alpha$ or $\alpha \in \beta$. That is, every nonempty set (or class) of ordinals has an $\in$-least element.*

Proof. Since C is not empty, we can fix an ordinal $z \in C$. We split into two cases. First, suppose that $z \cap C = \emptyset$. In this case, we let $\alpha = z$. To verify that α satisfies the lemma, fix an arbitrary $\beta \in C$. Since $\alpha \cap C = \emptyset$, we know $\beta \notin \alpha$. By Lemma 4.13, the only remaining possibilities are $\beta = \alpha$ or $\alpha \in \beta$. This completes the first case.

For the second case, assume that $z \cap C \neq \emptyset$. Let $B = z \cap C$ and notice that B is a nonempty subset of the ordinal z. Therefore, there is an $\alpha \in B$ such that $\alpha \cap B = \emptyset$. Notice that since $\alpha \in B$ and $B \subseteq z$, we have that $\alpha \in z$.

We claim that α satisfies the lemma. To see why, fix an arbitrary $\beta \in C$. If $\beta \in B$, then because $\alpha \cap B = \emptyset$, we know $\beta \notin \alpha$ and hence by Lemma 4.13 either $\alpha = \beta$ or $\alpha \in \beta$. On the other hand, if $\beta \notin B$, then $\beta \notin z$ (because $\beta \in C$). Hence, either $z = \beta$ (and hence $\alpha \in \beta$ because $\alpha \in z$) or $z \in \beta$ (and hence $\alpha \in \beta$ by Lemma 4.14 because $\alpha \in z$ and $z \in \beta$). □

At this point, we can draw some conclusions about the structure of the collection of all ordinals. Let $\mathbb{ON}$ denote the collection of all ordinals. We need to be a little careful because we will show that this collection is not a set. So, to be more accurate, consider $\alpha \in \mathbb{ON}$ as shorthand for the proposition "α is a transitive set such that the $\in$-relation defines a well order on α." (Formally, this collection forms a proper class in ZFC and we will return to exactly what that means later in the course.)

Lemma 4.16. *The $\in$-relation satisfies the axioms of a strict well order on $\mathbb{ON}$.*

Proof. On $\mathbb{ON}$, the $\in$-relation satisfies transitivity by Lemma 4.14, satisfies trichotomy by Lemma 4.13, satisfies irreflexivity by Lemma 4.11 and satisfies the condition to be a well order by Lemma 4.15. □

Because the $\in$-relation acts as a strict linear order on $\mathbb{N}$, we often write $\alpha < \beta$ in place of $\alpha \in \beta$ for $\alpha, \beta \in \mathbb{ON}$. We can also capture the nonstrict ordering relation on $\mathbb{ON}$.

Lemma 4.17. *The $\subseteq$-relation satisfies the axioms of a linear order on $\mathbb{ON}$.*

Proof. The $\subseteq$-relation is reflexive, antisymmetric and transitive on all sets. Therefore, we only need to show that it is linear on $\mathbb{ON}$. Fix $\alpha, \beta \in \mathbb{ON}$. We have to show that either $\alpha \subseteq \beta$ or $\beta \subseteq \alpha$. By Lemma 4.13, we know that either $\alpha \in \beta$, $\alpha = \beta$ or $\beta \in \alpha$. If $\alpha \in \beta$, then because β is a transitive set, we have $\alpha \subseteq \beta$ as required. Similarly, if $\beta \in \alpha$, then $\beta \subseteq \alpha$. Finally, if $\alpha = \beta$ then $\alpha \subseteq \beta$. □

Because the $\subseteq$-relation satisfies the axioms of a linear order on $\mathbb{ON}$, we often write $\alpha \leq \beta$ in place of $\alpha \subseteq \beta$ for $\alpha, \beta \in \mathbb{N}$. Notice that the strict ordering $<$ and the nonstrict ordering $\leq$ on $\mathbb{ON}$ are compatible in the sense that

$$\alpha \leq \beta \Leftrightarrow \alpha < \beta \text{ or } \alpha = \beta$$

(It is worth writing down these definitions and working out why this is true.)

Theorem 4.18. *The collection $\mathbb{ON}$ is not a set.*

Proof. Suppose for a contradiction that $\mathbb{ON}$ is set. We claim that the set $\mathbb{ON}$ is an ordinal. Because the $\in$-relation defines a strict well order on $\mathbb{ON}$, we only need to check that $\mathbb{ON}$ is transitive. Fix $\alpha \in \mathbb{ON}$ and we show $\alpha \subseteq \mathbb{ON}$. By Lemma 4.10, each $\beta \in \alpha$ is an ordinal and hence $\alpha \subseteq \mathbb{ON}$ as required. Therefore, $\mathbb{ON}$ is an ordinal. But, since $\mathbb{ON}$ is the set of all ordinals, we have $\mathbb{ON} \in \mathbb{ON}$ which contradicts Lemma 4.11. □

Although it might appear that the fact that $\mathbb{ON}$ is not a set would make it hard to work with $\mathbb{ON}$, this is in general not true. We do need to be careful at various points to remember that $\alpha \in \mathbb{ON}$ is really an abbreviation for "α is an ordinal" as opposed to a set collection process, but most of the time, we can safely work with $\mathbb{ON}$ as though it were a set.

The next three lemmas give us two ways to construct new ordinals from old ordinals.

Lemma 4.19. *Let A be a nonempty transitive set of ordinals. That is, if $x \in A$ and $x \subseteq A$. Then A is an ordinal.*

Proof. Fix a nonempty transitive set A of ordinals and we show that A is an ordinal. By assumption, A is a transitive set, so it remains to show that the $\in$-relation defines a strict well order on A. The $\in$-relation satisfies irreflexivity by Lemma 4.11, satisfies trichotomy by Lemma 4.13, satisfies transitivity by Lemma 4.14 and satisfies the property to be a well order by Lemma 4.15. □

Lemma 4.20. *Let B be a nonempty set of ordinals. Then*

$$\bigcup B = \bigcup_{\alpha \in B} \alpha$$

is an ordinal. This ordinal is typically denote by sup(B) and it is the least ordinal δ such that $\alpha \leq \delta$ for all $\alpha \in B$.

Proof. Fix a nonempty set B of ordinals and let $\sup(B) = \bigcup_{\alpha \in B} \alpha$. We would like to show that $\sup(B)$ is an ordinal using Lemma 4.19. To do so, we need to show that $\sup(B)$ is a nonempty transitive set of ordinals. However, notice that if $\sup(B)$ is empty, then $\sup(B) = 0$ and is an ordinal as required. Therefore, we can assume $\sup(B)$ is not empty.

The set $\sup(B)$ is a set of ordinals because if $\beta \in \sup(B)$, then $\beta \in \alpha$ for some $\alpha \in B$. Since α is an ordinal and $\beta \in \alpha$, β is an ordinal by Lemma 4.10.

To see that $\sup(B)$ is transitive, fix $\beta \in \sup(B)$ and $\gamma \in \beta$. We need to show that $\gamma \in \sup(B)$. By definition, $\beta \in \alpha$ for some $\alpha \in B$. Since α is an ordinal, $\beta \in \alpha$ implies $\beta \subseteq \alpha$ and hence $\gamma \in \alpha$. Therefore, $\gamma \in \alpha$ for some $\alpha \in B$ and hence $\gamma \in \sup(B)$. This completes the proof that $\sup(B)$ is transitive. Therefore, $\sup(B)$ is a nonempty transitive set of ordinals and hence by Lemma 4.19, $\sup(B)$ is an ordinal.

Next, we need to show that $\sup(B) = \delta$ where δ is the least ordinal such that $\alpha \leq \delta$ for all $\alpha \in B$. To do this, we first claim that if $\beta \in B$, then $\beta \leq \sup(B)$. To see why, fix $\beta \in B$. Then $\beta \subseteq \bigcup_{\alpha \in B} \alpha$ because β is one of the sets being collected in the union. Therefore, $\beta \subseteq \sup(B)$ which means $\beta \leq \sup(B)$.

Let C be the collection of ordinals such that $\gamma \in C$ if and only if $\alpha \leq \gamma$ for every $\alpha \in B$. By the previous paragraph, $\sup(B) \in C$ so C is nonempty. (To be clear, C is actually a

proper class.) By Lemma 4.15, C has an $\in$-least element δ. That is, there is a least ordinal δ such that $\alpha \leq \delta$ for all $\alpha \in B$. By the previous paragraph, we know that $\delta \leq \sup(B)$.

To show that $\sup(B) \leq \delta$, it suffices to show that $\bigcup_{\alpha \in B} \alpha \subseteq \delta$. Fix $\beta \in \bigcup_{\alpha \in B} \alpha$, and we show that $\beta \in \delta$. Since $\beta \in \bigcup_{\alpha \in B} \alpha$, we can fix an $\alpha \in B$ such that $\beta \in \alpha$. Since $\alpha \in B$, the definition of δ implies that $\alpha \leq \delta$ and hence $\alpha \subseteq \delta$. But, $\beta \in \alpha$ and $\alpha \subseteq \delta$ gives us $\beta \in \delta$ as required. □

Lemma 4.21. *Let α be an ordinal. Then $S(\alpha) = \alpha \cup \{\alpha\}$ is also an ordinal. Furthermore, for all ordinals β, $\beta \in S(\alpha)$ if and only if $\beta = \alpha$ or $\beta \in \alpha$.*

Proof. Since $x \in S(\alpha)$ if and only if $x \in \alpha$ or $x = \alpha$, $S(\alpha)$ is a set of ordinals. We claim $S(\alpha)$ is transitive. Fix $\beta \in S(\alpha)$ and $\gamma \in \beta$. We split into two cases. First, if $\beta \in \alpha$, then $\gamma \in \alpha$ because α is an ordinal, and hence $\gamma \in S(\alpha)$. Second, if $\beta \notin \alpha$, then the fact that $\beta \in S(\alpha)$ implies that $\beta = \alpha$. Therefore, $\gamma \in \beta$ implies that $\gamma \in \alpha$, and hence $\gamma \in S(\alpha)$. □

The ordinal $S(\alpha)$ is called the *successor of* α and we refer to the operation $S(x)$ mapping α to $S(\alpha)$ as the *successor operation*.

You should think of the last two lemmas as ordinal construction methods. First, if we have an ordinal α, we can construct a longer ordinal $S(\alpha)$ by adding a new greatest element. This is exactly the process that took us from the finite ordinal n to the finite ordinal $n+1$. The ordinal $S(\alpha)$ is often written as $\alpha + 1$ because it consists of the ordinal α with one new element placed on the end. We will see the notion of addition of ordinals defined formally later, although we will sneak in some of the notation now.

Second, if we have a set of ordinals, we can form a new ordinal which at least as long as each of the ordinals in this set. This is exactly the process that took us from the set of finite ordinals to the ordinal ω. Iterating these procedures generates many new (countable) ordinals. For example, starting with ω and applying the successor operation, we can generate

$$\begin{aligned} \omega : &\quad 0 < 1 < 2 < \cdots \\ \omega + 1 : &\quad 0 < 1 < 2 < \cdots < \omega \\ \omega + 2 : &\quad 0 < 1 < 2 < \cdots < \omega < \omega + 1 \\ \omega + 3 : &\quad 0 < 1 < 2 < \cdots < \omega < \omega + 1 < \omega + 2 \end{aligned}$$

Collecting the set of ordinal of the form $\omega + n$ for $n \in \mathbb{N}$ using the sup operation gives us

$$\begin{gathered} \omega + \omega = \sup(\{\omega + n \mid n \in \mathbb{N}\}) \\ 0 < 1 < 2 < \cdots < \omega < \omega + 1 < \omega + 2 < \omega + 3 < \cdots \end{gathered}$$

Of course, we can now start applying the successor operation again to $\omega + \omega$ to obtain a longer ordinal

$$\omega + \omega + 1 : \quad 0 < 1 < \cdots < \omega < \omega + 1 < \cdots \omega + \omega$$

and so on.

Theorem 4.22 (Fundamental Theorem of Ordinal Numbers)**.** *If W is a well order, then there is a unique ordinal α such that $W \cong \alpha$.*

Proof. Let $(W, \leq_W)$ be a well order. To see that the ordinal α is unique, suppose that $W \cong \alpha$ and $W \cong \beta$. Then, $\alpha \cong \beta$ and hence by Lemma 4.13, $\alpha = \beta$.

It remains to show that existence of α. If W is empty, then we let $\alpha = 0$ and note that $W \cong \alpha$ because both are empty. Therefore, we can assume that W is nonempty. Define

$$B = \{w \in W \mid I_W(w) \cong \beta \text{ for some ordinal } \beta\}.$$

Let f be the function with domain B such that for all $w \in B$, $f(w) = \beta$ for the unique ordinal β such that $I_W(w) \cong \beta$. Let A be the range of f.

First, we claim that A is a transitive set of ordinals. By definition, A is a set of ordinals. To see that it is transitive, fix $\beta \in A$ and $\gamma \in \beta$. Let $w \in W$ be such that $I_W(w) \cong \beta$ and let $f : \beta \to I_W(w)$ be the isomorphism. Let $w' \in W$ be such that $f(\gamma) = w'$. By Lemma 3.28, the restricted map $f \restriction I_\beta(\gamma)$ is an isomorphism from $I_\beta(\gamma)$ to $I_W(w')$. But, $I_\beta(\gamma) = \gamma$ by Lemma 4.10. So, $f \restriction I_\beta(\gamma)$ is an isomorphism from γ to $I_W(w')$. Therefore, $w' \in B$ and $\gamma \in A$. This completes the proof that A is transitive set of ordinals.

We would like use Lemma 4.19 to conclude that A is an ordinal. To do so, we need to know that A is nonempty. Since W is nonempty, W has a $\leq_W$-least element w. For this element w, we have $I_W(w) = \emptyset$ and hence $I_W(w) \cong 0$. Therefore, $w \in B$ and $0 \in A$. Having shown that A is nonempty, we conclude from Lemma 4.19 that A is an ordinal. Let $\alpha = A$.

Second, we claim that $f : B \to \alpha$ is an isomorphism. By the definition of α, this map is onto. Suppose for a contradiction that it is not one-to-one. Then there are $w' \neq w$ in B such that $f(w') = f(w)$. But, then $I_W(w') \cong I_W(w)$ with $w \neq w'$ which contradicts Corollary 3.26. Therefore, f is a bijection. Finally, to see that f is order preserving, fix $w' <_W w$ in B and fix β such that $f(w) = \beta$. Then, $\beta \cong I_W(w)$ and since $w' \in I_W(w)$, there is a $\gamma \in \beta$ such that $\gamma \cong I_W(w')$ using the appropriate restriction map. By the uniqueness of γ (which we established at the beginning of this proof), $f(w') = \gamma \in \beta$ as required.

Third, we claim that B is an initial segment of W. Fix $b \in B$ and $w \in W$ such that $w <_W b$. We need to show that $w \in B$. Since $b \in B$, we can fix an ordinal β and an isomorphism $g : \beta \to I_W(b)$. Since $w \in I_W(b)$, there is a $\gamma \in \beta$ such that $g(\gamma) = w$. The restriction $g \restriction \gamma$ gives an isomorphism $\gamma \cong I_W(w)$ witnessing that $w \in B$.

Finally, we claim that $B = W$. For a contradiction, suppose that $W \setminus B \neq \emptyset$ and hence B is a proper initial segment of W. Fix $w \in W$ such that $B = I_W(w)$. Define $g : B \cup \{w\} \to S(\alpha)$ by $g(b) = f(b)$ for all $b \in B$ and $g(w) = \alpha$. Because $f : B \to \alpha$ is a bijection, g is also a bijection. (It maps the single element w not in the domain of f to the single element α not in the range of f.) Note that g is order preserving because w is the $<_W$-greatest element of $B \cup \{w\}$ and α is the $\in$-greatest element of $S(\alpha)$. Therefore, g is an isomorphism witnessing $w \in B$. This contradiction the fact that $w \in W \setminus B$. □

Definition 4.23. For a well order W, we refer to the unique ordinal isomorphic to W as the *order type of* W and denote it by $\mathrm{Otp}(W, \leq_W)$ or just $\mathrm{Otp}(W)$. Thus, $W \cong \mathrm{Otp}(W)$.

We end this section with a description of the basic arithmetic operations on $\mathbb{ON}$. Consider the successor operation. Given an ordinal α, we defined $S(\alpha) = \alpha \cup \{\alpha\}$. Therefore, we can think of $S(x)$ as a function $S : \mathbb{ON} \to \mathbb{ON}$. Notice that $\alpha < S(\alpha)$ and that $S(\alpha)$ is actually the successor of α in the ordering on $\mathbb{ON}$. That is, there cannot be an ordinal β such that

$\alpha < \beta < S(\alpha)$. To see why, suppose there is such a β. Since $\alpha < \beta$, we have $\alpha \subsetneq \beta$ and hence β must contain something not in α. However, $\beta < S(\alpha) = \alpha \cup \{\alpha\}$ implies that the only thing β could contain that is not in α is α itself. That is, we must have $\alpha \in \beta$. But, then

$$\alpha \cup \{\alpha\} \subseteq \beta \subseteq S(\alpha) = \alpha \cup \{\alpha\}$$

and hence $\beta = S(\alpha)$ which contradictions $\beta < S(\alpha)$.

The successor function gives us a useful way to partition $\mathbb{ON}$ into those ordinals in the range of $S(x)$ and those ordinals not in the range of $S(x)$.

Definition 4.24. An ordinal α is called a *successor ordinal* if there is an ordinal $\beta < \alpha$ such that $S(\beta) = \alpha$. An ordinal α is a *limit ordinal* if it is not a successor ordinal.

The ordinal 0 counts as a limit ordinal since it is the least ordinal and hence not the successor of anything. The first infinite limit ordinal is ω.

Consider addition on $\mathbb{ON}$. We will explain how to define addition in two different ways. For the first method, fix $\alpha, \beta \in \mathbb{ON}$. Because α and β are well orders, we can add them as described in the previous section. Let $W = \alpha \times \{0\} \cup \beta \times \{1\}$ and define a (strict) well order on W by

$$\langle u, v\rangle \prec \langle x, y\rangle \Leftrightarrow v <_{\mathbb{N}} y \text{ or } \big(v = y \text{ and } u \in x\big)$$

As described in the previous section, $(W, \prec)$ puts down a copy of α and then puts down a copy of β so that all the elements of β come after the elements of α. We define the ordinal sum $\alpha + \beta$ by

$$\alpha + \beta = \mathrm{Otp}(W).$$

We can also define $\alpha + \beta$ by transfinite recursion. Recall that to define a function f by recursion on $\mathbb{N}$, we specify the value of $f(0)$ and we specify the value of $f(n+1)$ assuming we have already given values for $f(0), \ldots, f(n)$.

Example 4.25. To define $f(n) = 2^n$ on $\mathbb{N}$ by recursion we set $f(0) = 1$ and $f(n+1) = 2 \cdot f(n)$. Notice that for this recursive definition, we already need to have define the multiplication function (or at least the doubling function). Also, if we want to check that our recursive definition is correct, we would proceed by induction on $\mathbb{N}$. For the base case, $f(0) = 1$ and $2^0 = 1$, so $f(0) = 2^0$. For the induction case, assume that $f(n) = 2^n$. Then, $f(n+1) = 2 \cdot f(n)$ by definition. But, $2f(n) = 2 \cdot 2^n$ by the induction hypothesis and $2 \cdot 2^n = 2^{n+1}$. Therefore, our recursive definition defines the correct function.

We can use recursion on $\mathbb{N}$ to define functions of two variables $f(n, m)$ by fixing the first variable and giving a recursion definition on the second variable. Formally, for each n, we recursively define a function $g_n(m)$. This process gives us a family of functions $g_n(m)$ parameterized by $n \in \mathbb{N}$. We collect them together (by taking a union) into a single function $f(n, m)$.

Example 4.26. To define $f(n, m) = n + m$ on $\mathbb{N}$ by recursion, we fix the value of n and proceed by recursion on m. We set $f(n, 0) = n$ and $f(n, m+1) = f(n, m) + 1$. Again, notice that for this recursive definition, we need to have defined the successor function already.

We can define a function $f : \mathbb{ON} \to \mathbb{ON}$ recursively in much the same way. To do this formally, we would need to give an appropriate theorem and proof, which we will not do in this course. (You can find this information in an introductory set theory book and if there is sufficient interest, I am happy to go through it with you.) However, it is worth noting that the function f would really be a proper class object as opposed to a set of pairs.

The main difference between defining $f : \mathbb{N} \to \mathbb{N}$ by recursion and defining $f : \mathbb{ON} \to \mathbb{ON}$ by recursion is that in the $\mathbb{ON}$ case, we have to deal with infinite limit ordinals. That is, there are typically three cases in the definition of $f(\beta)$: specifying $f(0)$, specifying $f(S(\beta))$ assuming you know $f(\beta)$, and specifying $f(\beta)$ when β is an infinite limit ordinal assuming you know $f(\gamma)$ for all $\gamma < \beta$. (In the successor case, you can assume you know $f(\gamma)$ for all $\gamma < S(\beta)$ but often you only use the value in $f(\beta)$ in the recursive definition.)

We illustrate this process by defining $\alpha + \beta$ by recursion on β. As in the example of definition addition on $\mathbb{N}$, we fix α and set

$$\begin{gathered}\alpha + 0 = \alpha \\ \alpha + S(\beta) = S(\alpha + \beta) \text{ for the successor case} \\ \alpha + \beta = \sup(\{\alpha + \gamma \mid \gamma < \beta\}) \text{ for a limit } \beta\end{gathered}$$

Given that we now have two definitions of ordinal addition, we could check that they are the same by transfinite induction. Although we won't do it in these notes, it is a good exercise to go through.

Using ordinal addition, we can define ordinal multiplication $\alpha \cdot \beta$ by transfinite recursion on β.

$$\begin{gathered}\alpha \cdot 0 = 0 \\ \alpha \cdot S(\beta) = \alpha \cdot \beta + \alpha \\ \alpha \cdot \beta = \sup(\{\alpha \cdot \gamma \mid \gamma < \beta\}) \text{ for limit } \beta\end{gathered}$$

Then, using ordinal multiplication, we can define ordinal exponentiation α^β by transfinite recursion on β.

$$\begin{gathered}\alpha^0 = 1 \\ \alpha^{S(\beta)} = \alpha^\beta \cdot \alpha \\ \alpha^\beta = \sup(\{\alpha^\gamma \mid \gamma < \beta\}) \text{ for limit } \beta\end{gathered}$$

5 Cardinal numbers

In the second section of these notes, we showed how to compare the relative sizes of two sets. That is, given two sets A and B, we defined $|A| = |B|$ and $|A| \leq |B|$. However, this notion of relative size only allows us to compare the sizes of two sets. For finite sets, we know how to do better than this. If a set A is finite, then we can specify uniquely what its size is. That is, we can say that A has exactly n elements for some $n \in \mathbb{N}$. Notice that in this context we are thinking of n in a cardinal manner (i.e. as a size indicator) as opposed to in an ordinal

manner (i.e. as a queue length indicator). When n is used as a size indicator, we call it a *finite cardinal number* and when n is used as a queue length indicator, we call it a *finite ordinal number*.

In this section, we want to extend this idea of uniquely specifying the size of a set to handle the case when the set is infinite. That is, we want to extend the natural numbers as size indicators (or cardinals) so that we can uniquely assign a cardinal number to each set indicating its size. We will use the ordinal numbers to do this.

To give the idea, consider a set $A = \{a, b, c\}$. Notice that there are several ways to define a linear order on the elements of A. For example, we might set $a < b < c$. Or, we might set $b < a < c$, and so on. You might notice that each of these linear orders on A has the same well order type. As long as A is finite, that will be true. However, it fails badly when A becomes infinite. For example, consider $\mathbb{N}$. We might order $\mathbb{N}$ in the usual way $0 < 1 < 2 < \cdots$. However, we might order $\mathbb{N}$ in any of the following ways

$$1 < 2 < 3 < \cdots < 0$$
$$2 < 3 < 4 < \cdots < 0 < 1$$
$$0 < 2 < 4 < \cdots < 1 < 3 < 5 < \cdots$$

We can put orders on $\mathbb{N}$ which are not well orders at all. For example, $0 > 1 > 2 > \cdots$. In fact, let $(L, \leq_L)$ be any countable linear order and fix a bijection $f : \mathbb{N} \to L$ (which we can do because L is countable). Define a linear order $\leq'$ on $\mathbb{N}$ by $n \leq' m$ if and only if $f(n) \leq_L f(m)$. By the definition of $\leq'$, the bijection f is an isomorphism between $(\mathbb{N}, \leq')$ and $(L, \leq_L)$. Therefore, we can order $\mathbb{N}$ in any countable linear order type whatsoever!

The idea for uniquely specifying the size of A is to consider all the ways in which A can be well ordered. If A is infinite, there will be many different ways to define a well ordering of A. However, each of these well orders corresponds to a unique ordinal. The set of ordinals obtained from the well ordering of A must have a least element. We use this least element to uniquely specify the size of A. To do this formally, we first need to know that for any set A, we can define a well ordering on A.

Lemma 5.1 (Well Ordering Principle)**.** *Let A be a set. There is a binary relation $\leq_A$ on A such that $(A, \leq_A)$ is well order.*

The Well Ordering Principle is a version of the Axiom of Choice. More specfically, over ZF it is equivalent to the Axioms of Choice. For now, we will put off giving a proof of this equivalence and we will possibly return to it at the end of the course.

Using the Well Ordering Principle, we can assign to each set an ordinal which represents the size (or cardinality) of the set. We assign this ordinal as follows. Fix a set A. By the Well Ordering Principle, the follow set is nonempty.

$$\mathcal{O}(A) = \{\alpha \in \mathbb{ON} \mid \alpha \cong (A, \leq_A) \text{ for some well ordering } \leq_A \text{ of } A\}$$

Since $\mathcal{O}(A)$ is a nonempty set of ordinals, it contains a least element. We define the *cardinality of* A by

$$\text{Card}(A) = \text{ the least element of } \mathcal{O}(A).$$

Thus, Card(A) is an ordinal α which represents the shortest possible way to put the elements of A is a well ordered list. Notice that there is a bijection between A and Card(A). Namely, fix a well ordering $\leq_A$ such that $(A, \leq_A) \cong \text{Card}(A)$ and an isomorphism $f : (A, \leq_A) \to \text{Card}(A)$. Then f is a bijection between A and Card(A), so $|A| = |\text{Card}(A)|$.

Example 5.2. Fix $n \in \mathbb{N}$ with $n > 0$. Let $A = \{a_0, a_1, \ldots, a_{n-1}\}$ be a set with n elements. Every well ordering of A is isomorphic to the ordinal n and therefore $\mathcal{O}(A) = \{n\}$. Since the least (in fact only) element of $\mathcal{O}(A)$ is n, we have $\text{Card}(A) = n$.

Example 5.3. Consider Card$(\emptyset)$. Because $\emptyset$ has no elements, the empty relation is a well ordering of it. This well ordering is isomorphic to the ordinal 0. Therefore, $\mathcal{O}(\emptyset) = \{0\}$ and $\text{Card}(\emptyset) = 0$.

Example 5.4. Consider Card$(\mathbb{N})$. We have already shown that for every countable linear order $(L, \leq_L)$, there is a linear order $\leq'$ on $\mathbb{N}$ such that $(\mathbb{N}, \leq') \cong (L, \leq_L)$. In particular, this fact is true for every countable well order and hence for every countable ordinal. Therefore, for every countable α, there is a well order $\leq'$ of $\mathbb{N}$ such that $(\mathbb{N}, \leq') \cong \alpha$. Since every well ordering of $\mathbb{N}$ is isomorphic to some countable ordinal, we have show that

$$\mathcal{O}(\mathbb{N}) = \{\alpha \in \mathbb{ON} \mid \alpha \text{ is countable}\}.$$

The least ordinal in this set is ω and hence $\text{Card}(\mathbb{N}) = \omega$.

Because Card(A) is an ordinal, we have a new way to compare the size of two sets. Given sets A and B, we can compare Card(A) and Card(B). The next lemma shows that this method of comparing the size of two sets corresponds exactly to the method of looking for bijections between sets.

Lemma 5.5. *For any sets A and B, $|A| = |B|$ if and only if Card(A) = Card(B). Furthermore, $|A| \leq |B|$ if and only if Card$(A) \leq$ Card(B).*

Proof. First, suppose that $\text{Card}(A) = \text{Card}(B) = \alpha$. As noted above, there are bijections between A and α and between B and α. Therefore, there is also a bijection between A and B.

Similarly, suppose $\text{Card}(A) = \alpha$, $\text{Card}(B) = \beta$ and $\alpha \leq \beta$. Fix bijections $f : A \to \alpha$ and $g : \beta \to B$. Since $\alpha \leq \beta$, there is an injective map $i : \alpha \to \beta$. Composing these maps as $g \circ i \circ f : A \to B$ gives an injection from A into B.

Second, suppose that $|A| = |B|$ and fix a bijection $h : A \to B$. Let $\text{Card}(A) = \alpha$ and $\text{Card}(B) = \beta$. We need to show that $\alpha = \beta$. Fix a well order $\leq_B$ and an isomorphism $f : (B, \leq_B) \to \beta$. Define a well order $\leq_A$ on A by $x \leq_A y$ if and only if $h(x) \leq_B h(y)$. You should check that $\leq_A$ is a well order and that h gives an isomorphism between $(A, \leq_A)$ and $(B, \leq_B)$. Since $(A, \leq_A) \cong (B, \leq_B) \cong \beta$, we have $\beta \in \mathcal{O}(A)$. Therefore, since α is the least element of $\mathcal{O}(A)$, we have $\alpha \leq \beta$.

Switching the roles of A and B, you can use essentially the same argument to show that $\beta \leq \alpha$ and hence $\alpha = \beta$ as required. □

Because of this connection between $\text{Card}(A)$ and $|A|$, people typically equate $|A|$ and $\text{Card}(A)$. That is, given our definitions, the notation $|A|$ by itself is not defined, but we define it by $|A| = \text{Card}(A)$. Given this definition, the statement $|A| = |B|$ becomes ambiguous. It can mean that there is a bijection between A and B or that $\text{Card}(A) = \text{Card}(B)$. However, by Lemma 5.5, these concepts are equivalent, so the ambiguity causes no problems. We will use $|A|$ instead of $\text{Card}(A)$ because it is shorter.

Definition 5.6. An ordinal number α is called a *cardinal number* if there is a set A such that $|A| = \alpha$.

By Examples 5.2 and 5.3, we know that each $n \in \omega$ is cardinal, and by Example 5.4, we know that ω is a cardinal. We next give a useful characterization for when an ordinal number is a cardinal.

Lemma 5.7. *For an ordinal α, the following are equivalent.*

(1) α *is a cardinal*

(2) $|\alpha| = \alpha$

(3) $\beta < \alpha$ *implies* $\beta < |\alpha|$

(4) $\beta < \alpha$ *implies* $|\beta| < |\alpha|$

(5) $\beta < \alpha$ *implies* $|\beta| \neq |\alpha|$

Proof. To see (1) implies (2), fix a set A such that $|A| = \alpha$. Because there is a bijection between A and α, we have $|A| = |\alpha|$. Therefore, $|\alpha| = |A| = \alpha$ as required.

To see (2) implies (3), fix $\beta < \alpha$. By (2), $\alpha = |\alpha|$, so $\beta < \alpha = |\alpha|$ giving $\beta < |\alpha|$ as required.

To see that (3) implies (4), notice that the identity function $\text{id} : \beta \to \beta$ is an isomorphism, so we have $|\beta| \leq \beta$. Therefore, if $\beta < \alpha$, then $\beta < |\alpha|$ by (3). Hence $|\beta| \leq \beta < |\alpha|$ which means $|\beta| < |\alpha|$ as required.

To see that (4) implies (5), fix $\beta < \alpha$. By (4), $|\beta| < |\alpha|$ and hence $|\beta| \neq |\alpha|$ are required.

To see that (5) implies (1), assume that (5) holds and consider what $|\alpha|$ could be. As noted in the proof of (3) implies (4), we know that $|\alpha| \leq \alpha$. However, if $|\alpha| < \alpha$, then $|\alpha| = \beta$ for some $\beta < \alpha$. In particular, there is a bijection between α and β which means $|\alpha| = |\beta|$ contradicting (5). □

Cardinal numbers are often denoted by the Greek letters κ, λ and μ, while ordinals are often denoted by α, β and γ. However, since cardinal numbers are ordinals, it is important to keep in mind that even when κ is explicitly defined as a cardinal, it can be treated as an ordinal. Conversely, since some ordinals are cardinals, an ordinal α might be a cardinal as well (or might switch to being viewed as a cardinal in the middle of a proof). The next lemma is an example of this phenomenon.

Lemma 5.8. *If κ is an infinite cardinal, then κ is a limit ordinal.*

Proof. If κ is an infinite cardinal, then (as an ordinal), $\kappa \geq \omega$. Suppose for a contradiction that $\kappa = S(\alpha)$. It suffices to define a bijection $f : \kappa \to \alpha$ showing $|\kappa| < \kappa$ which contradicts Lemma 5.7.

Since $\kappa = S(\alpha) = \alpha \cup \{\alpha\}$ and $\omega \leq \kappa$, an element $\beta \in \kappa$ satisfies exactly one of the following conditions: $\beta \in \omega$ or $\omega \leq \beta < \alpha$ or $\beta = \alpha$. Define f by

$$f(\beta) = \begin{cases} 0 & \text{if } \beta = \alpha \\ \beta + 1 & \text{if } \beta \in \omega \\ \beta & \text{if } \omega \leq \beta < \alpha \end{cases}$$

The definition of f is like the solution to Hilbert's Hotel. We move each element of the initial segment ω of κ over by one position. This movement opens up a spot for α (the greatest element of κ) to move to 0. Then, we leave all the remaining elements of κ as they were. You should check that f is a bijection to complete the proof. □

We can use the fact that every cardinal is an ordinal in a second useful way. Fix a cardinal κ. Let C_κ be the class of cardinals strictly greater than κ. C_κ is nonempty because $|\kappa| < |\mathcal{P}(\kappa)|$ by Cantor's Theorem. In fact, C_κ is a proper class. Because C_κ is a nonempty class of ordinals, it has a least element and we denote this least element by κ^+. That is, κ^+ is the least cardinal greater than κ.

To distinguish ω as a cardinal from ω as an ordinal, set theorists often use $\aleph_0$ to denote ω as a cardinal. That is, $\aleph_0$ is a notation for ω when it is explicitly being used as a cardinal. However, you should be careful because ω is often used for this cardinal as is ω_0.

Using the operation $\kappa \mapsto \kappa^+$, we can define an initial segment of the class of cardinals.

$$\aleph_0 = \omega = \text{the least infinite cardinal}$$
$$\aleph_1 = \aleph_0^+ \qquad \aleph_2 = \aleph_1^+ \qquad \aleph_3 = \aleph_2^+ \qquad \text{and so on}$$

In fact, we can continue defining this sequence of cardinals transfinitely. We define a function $\alpha \mapsto \aleph_\alpha$ by transfinite recursion on $\mathbb{ON}$ as follows.

$$\aleph_0 = \omega$$
$$\aleph_{\alpha+1} = \aleph_\alpha^+$$
$$\aleph_\alpha = \sup(\{\aleph_\gamma \mid \gamma < \alpha\}) \text{ for limit } \alpha$$

We need to check that if α is an infinite limit ordinal, then $\aleph_\alpha$ is actually a cardinal. Suppose for a contradiction that $\aleph_\alpha$ is not a cardinal. By Lemma 5.7, $|\aleph_\alpha| = \beta$ for some $\beta < \aleph_\alpha$. Because $\aleph_\alpha$ is defined as a supremum, $\beta < \aleph_\alpha$ implies that $\beta \leq \aleph_\gamma$ for some $\gamma < \alpha$. However, α is a limit ordinal, so $\gamma + 1 < \alpha$. Since $\aleph_{\gamma+1} = \aleph_\gamma^+$, we know $\aleph_\gamma < \aleph_{\gamma+1}$. Therefore, we have shown

$$|\aleph_\alpha| = \beta < \aleph_{\gamma+1} \leq \aleph_\alpha$$

Since $\aleph_{\gamma+1} \leq \aleph_\alpha$, we have $\aleph_{\gamma+1} \subseteq \aleph_\alpha$ and hence $|\aleph_{\gamma+1}| \leq |\aleph_\alpha|$. Since $|\aleph_\alpha| = \beta$, we have $|\aleph_{\gamma+1}| \leq \beta$. On the other hand, $\aleph_{\gamma+1}$ is a cardinal, so $|\aleph_{\gamma+1}| = \aleph_{\gamma+1}$. Since $\beta < \aleph_{\gamma+1}$, we have $\beta < |\aleph_{\gamma+1}|$. Therefore, we have show $|\aleph_{\gamma+1}| \leq \beta$ and $\beta < |\aleph_{\gamma+1}|$ for the desired contradiction.

Putting this information together, we see that the collection of cardinal numbers is well ordered and indexed by the ordinals. The aleph-notation makes this connection explicit and give us an ordinal indexed list of all cardinals.

$$\aleph_0 < \aleph_1 < \cdots < \aleph_\omega < \aleph_{\omega+1} < \cdots < \aleph_{\omega+\omega} < \cdots$$

We would like to define arithmetic operations on the cardinal numbers. Before giving the formal definitions, consider how we define arithmetic operations on the natural numbers when they are viewed in a cardinal manner. We can think of $n+m$ as representing the total number of objects when we combine a set of n many objects with a disjoint set of m many objects. That is, if $|A| = n$ and $|B| = m$, the $n + m = |A \cup B|$ as long as A and B are disjoint.

We run into one subtlety formalizing this notion. Since the cardinal numbers n and m are really ordinals, they are sets. However, they are not disjoint sets because $n = \{0, 1, \ldots, n-1\}$ and $m = \{0, 1, \ldots, m-1\}$. Even worse, either $n \subseteq m$ or $m \subseteq n$! Therefore, we would like to have a canonical way to fix disjoint sets of size n and m. We encountered this problem before in defining ordinal addition $\alpha + \beta$ when we wanted to place a queue of length β after a queue of length α. We solved the problem in the context of ordinal addition by considering $\alpha \times \{0\}$ and $\beta \times \{1\}$. We can solve our current problem in exactly the same way.

Definition 5.9. Let κ and λ be cardinals. We define *cardinal addition* by

$$\kappa \oplus \lambda = |(\kappa \times \{0\}) \cup (\lambda \times \{1\})|.$$

That is, $\kappa \oplus \lambda$ is the cardinality of the disjoint union of a set of size κ and a set of size λ.

For the finite cardinals, $n \oplus m$ is the same as the usual addition $n +_{\mathbb{N}} m$ on $\mathbb{N}$ and is also the same as the ordinal addition $n + m$. However, when we deal with infinite cardinals, cardinal addition is not the same as ordinal addition.

Example 5.10. Since the union of two countable is countable, we have $\omega \oplus \omega = \omega$. However, in terms of ordinal addition, $\omega + \omega \neq \omega$.

Unlike ordinal addition, cardinal addition is commutative. That is, $\kappa \oplus \lambda = \lambda \oplus \kappa$ because there is a bijection between $(\kappa \times \{0\}) \cup (\lambda \times \{1\})$ and $(\lambda \times \{0\}) \cup (\kappa \times \{1\})$ given by toggling the second component between 0 and 1.

To motivate cardinal multiplication, we consider a finite example. Think about 2×3 as taking 2 disjoint sets with 3 elements each and counting the total number of elements. We can think of putting these objects in a 2 by 3 array and counting the total number of objects in the array. We follow the same outline for general cardinal multiplication.

Definition 5.11. We define *cardinal multiplication* by

$$\kappa \otimes \lambda = |\kappa \times \lambda|.$$

That is, $\kappa \otimes \lambda$ is the cardinality of the union of κ many disjoint sets of size λ.

Many of the same points about cardinal addition apply to cardinal multiplication. Also, since there is a bijection between $\kappa \times \lambda$ and $\lambda \times \kappa$ given by $f(\langle x, y\rangle) = \langle y, x\rangle$, we have $\kappa \otimes \lambda = \lambda \otimes \kappa$. For finite ordinals, $n \otimes m$ is the same as $n \cdot_{\mathbb{N}} m$ in $\mathbb{N}$ which is the same as $n \cdot m$ as ordinals. But for infinite cardinals, cardinal multiplication and ordinal multiplication will not be the same.

Example 5.12. Since $\omega \times \omega$ is countable, we have $\omega \otimes \omega = \omega$. However, as ordinals, $\omega \times \omega \neq \omega$.

Fortunately, there is an easy way to calculate $\kappa \oplus \lambda$ and $\kappa \otimes \lambda$ when at least one of κ and λ is infinite. To give this formula, we first need to prove that if κ is infinite, then the union of κ many sets of size κ has size κ, i.e. general the fact that a countable union of countable sets is countable to larger infinite cardinal sizes.

Theorem 5.13. *If κ is an infinite cardinal, then $\kappa \otimes \kappa = \kappa$.*

Proof. Fix an infinite cardinal. We show by induction that on α that

$$\omega \leq \alpha \leq \kappa \to |\alpha| \otimes |\alpha| = |\alpha| \tag{1}$$

Notice that when $\alpha = \omega$, this statement says that a countable union of countable sets is countable which we know already. To proceed by induction, we fix an ordinal λ such that $\omega < \lambda \leq \kappa$ and assume Equation (1) holds for all infinite $\alpha < \lambda$. We split into two case.

Our first case is when λ is not a cardinal. By Lemma 5.7, there is an $\alpha < \lambda$ such that $|\alpha| = |\lambda|$. We obtain Equation (1) for λ from

$$|\lambda| \otimes |\lambda| = |\alpha| \otimes |\alpha| = |\alpha| = |\lambda|$$

The first and third equalities follows from $|\lambda| = |\alpha|$ and the second equality follows from Equation (1) holding for α. This completes the first case.

The nontrivial case is when λ is a cardinal. By Lemma 5.7, $|\lambda| = \lambda$ and so

$$|\lambda| \otimes |\lambda| = \lambda \otimes \lambda = |\lambda \times \lambda|.$$

Therefore, to show that $|\lambda| \otimes |\lambda| = |\lambda|$, it suffices to show that $\lambda = |\lambda \times \lambda|$. The inequality $\lambda \leq |\lambda \times \lambda|$ follows immediately because the function sending $\alpha \mapsto \langle \alpha, 0\rangle$ is a one-to-one map from λ into $\lambda \times \lambda$. That is, this map shows that $|\lambda| \leq |\lambda \times \lambda|$ and since $|\lambda| = \lambda$, we have $\lambda \leq |\lambda \times \lambda|$.

What remains to show is that $|\lambda \times \lambda| \leq \lambda$. By definition, $|\lambda \times \lambda|$ is the least ordinal γ such that there is a well ordering of $\lambda \times \lambda$ which is isomorphic to γ, i.e. has order type γ. Therefore, to complete the proof, we need to show that there is a well ordering $\preceq$ of $\lambda \times \lambda$ such that $\mathrm{Otp}(\lambda \times \lambda, \preceq) \leq \lambda$. The rest of the proof consists of giving this well ordering $\preceq$.

The induction hypothesis is that $|\alpha| \otimes |\alpha| = |\alpha|$ for all $\alpha < \lambda$. Therefore, we know

$$|\alpha \times \alpha| = |\alpha| \otimes |\alpha| = |\alpha| \leq \alpha < \lambda \tag{2}$$

for all $\alpha < \lambda$. You should think about why $|\alpha \times \alpha| = |\alpha| \otimes |\alpha|$. This fact is true about every ordinal α. The induction hypothesis is just the fact that $|\alpha| \otimes |\alpha| = |\alpha|$. The important fact from Equation (2) we will use is that if $\alpha < \lambda$, then $|\alpha \times \alpha| < \lambda$.

We define a well order $\preceq$ on $\lambda \times \lambda$ as follows.

$$\langle \alpha, \beta \rangle \preceq \langle \gamma, \delta \rangle \Leftrightarrow \max\{\alpha, \beta\} < \max\{\gamma, \delta\} \text{ or } \big(\max\{\alpha, \beta\} = \max\{\gamma, \delta\} \text{ and } \langle \alpha, \beta \rangle \leq_{\text{lex}} \langle \gamma, \delta \rangle\big)$$

Recall that $\langle \alpha, \beta \rangle \leq_{\text{lex}} \langle \gamma, \delta \rangle$ if and only if either $\alpha < \gamma$ or both $\alpha = \gamma$ and $\beta \leq \delta$. You should check that $\preceq$ is a well order of $\lambda \times \lambda$.

Consider an element $\langle \gamma, \delta \rangle \in \lambda \times \lambda$ and the initial segment $I_{\preceq}(\langle \gamma, \delta \rangle)$ determined by $\langle \gamma, \delta \rangle$ in the well order $(\lambda \times \lambda, \preceq)$. We want to show that the order type of this initial segment is less than λ. Notice that if $\text{Otp}(I_{\preceq}(\langle \gamma, \delta \rangle)) < \lambda$ for every $\langle \gamma, \delta \rangle \in \lambda \times \lambda$, then $\text{Otp}(\lambda \times \lambda, \preceq) \leq \lambda$ and we are done.

Fix an element $\langle \gamma, \delta \rangle \in \lambda \times \lambda$ and let $\varepsilon = \max\{\gamma, \delta\} + 1$. We claim that $I_{\preceq}(\langle \gamma, \delta \rangle) \subseteq \varepsilon \times \varepsilon$. To see why, consider an element $\langle \alpha, \beta \rangle \in I_{\preceq}(\langle \gamma, \delta \rangle)$. Since $\langle \alpha, \beta \rangle \prec \langle \gamma, \delta \rangle$, we know $\max\{\alpha, \beta\} \leq \max\{\gamma, \delta\} < \varepsilon$. Hence $\alpha, \beta < \varepsilon$ and $\langle \alpha, \beta \rangle \in \varepsilon \times \varepsilon$ proving this claim.

We now know that $I_{\preceq}(\langle \gamma, \delta \rangle) \subseteq \varepsilon \times \varepsilon$. Therefore,

$$|I_{\preceq}(\langle \gamma, \delta \rangle)| \leq |\varepsilon \times \varepsilon|. \tag{3}$$

However, because λ is an infinite cardinal, it is a limit ordinal. Therefore, $\gamma, \delta < \lambda$ implies $\gamma + 1, \delta + 1 < \lambda$ and so $\varepsilon < \lambda$. This means that we can apply the induction hypothesis in Equation (2) to ε and conclude that $|\varepsilon \times \varepsilon| < \lambda$. Combining this fact with Equation (3), we have

$$|I_{\preceq}(\langle \gamma, \delta \rangle)| < \lambda. \tag{4}$$

Finally, by Lemma 5.7, $|I_{\preceq}(\langle \gamma, \delta \rangle)| < \lambda$ implies that $\text{Otp}(I_{\preceq}(\langle \gamma, \delta \rangle)) < \lambda$. □

Theorem 5.14. *Let κ and λ be cardinals such that $\kappa, \lambda \geq 2$ and at least one of κ and λ is infinite. Then $\kappa \oplus \lambda = \kappa \otimes \lambda = \max\{\kappa, \lambda\}$.*

Proof. Since cardinal addition and multiplication are commutative, we can assume without loss of generality that $\lambda \leq \kappa$. We show $\kappa \oplus \lambda \leq \kappa \otimes \lambda$ with the following calculation.

$$\begin{aligned} \kappa \oplus \lambda &= |(\kappa \times \{0\}) \cup (\lambda \times \{1\})| \\ &\leq |(\kappa \times \{0\}) \cup (\kappa \times \{1\})| \\ &= |\kappa \times 2| \\ &\leq |\kappa \times \lambda| \\ &= \kappa \otimes \lambda \end{aligned}$$

The first line is the definition of $\kappa \oplus \lambda$. The second line follows because $\lambda \leq \kappa$ and so $\lambda \times \{1\} \subseteq \kappa \times \{1\}$. The third follows because $2 = \{0, 1\}$ and hence $\kappa \times 2 = (\kappa \times \{0\}) \cup (\kappa \times \{1\})$. The fourth line follows because $2 \leq \lambda$ and the last line follows by the definition of $\kappa \otimes \lambda$.

Next, we show that $\kappa \otimes \lambda \leq \kappa$ with the following calculation.

$$\kappa \otimes \lambda = |\kappa \times \lambda| \leq |\kappa \times \kappa| = \kappa$$

The first equality is the definition of $\kappa \otimes \lambda$. The inequality follows because $\lambda \leq \kappa$ and hence $\kappa \times \lambda \subseteq \kappa \times \kappa$. The last equality follows from Theorem 5.13. Notice that κ is infinite because $\lambda \leq \kappa$ and at least one of λ and κ is infinite.

We can now combine these facts to get the desired conclusion.

$$\kappa \leq \kappa \oplus \lambda \leq \kappa \otimes \lambda \leq \kappa$$

□

We have completely determined the values for cardinal addition and multiplication. (You should think about this point and write out all the cases. There are a couple of trivial cases which we have not said anything about.)

The last arithmetic operation on the cardinal numbers is exponentiation. Viewing κ and λ as sets, κ^λ denotes the set of all functions from κ to λ. We define the cardinal exponential to be the size of this set of functions. Unfortunately, we use the same notation for the cardinal exponential and the set of functions from λ to κ and rely on the context to tell us which is intended.

Definition 5.15. Let κ and λ be cardinals. The *cardinal exponentiation* κ^λ is defined by

$$\kappa^\lambda = |\kappa^\lambda|$$

Example 5.16. If $n, m \in \omega$, then n^m (in cardinal exponentiation) denotes the number of functions from a set of size n to a set of size m. This is exactly the same as n^m in the natural numbers.

Example 5.17. If κ is an infinite cardinal, then $\kappa^2 = \kappa$. To see why, it helps to work through the overloaded notation. The notation κ^2 has (at least) three meanings at this point. First, we can view κ^2 as the set of ordered pairs of elements from κ.

$$\kappa^2 = \{\langle \alpha, \beta \rangle \mid \alpha, \beta \in \kappa\}$$

From this point of view, $\kappa^2 = \kappa \times \kappa$ (the Cartesian product) and $|\kappa^2| = |\kappa \times \kappa| = \kappa \otimes \kappa = \kappa$ as long as κ is infinite.

Second, we can view κ^2 as the set of functions from 2 into κ. There is a natural bijection between κ^2 as a set of functions and κ^2 as a set of ordered pairs. To an ordered pair $\langle \alpha, \beta \rangle \in \kappa^2$, we associate the function $f_{\langle \alpha, \beta \rangle} \in \kappa^2$ given by $f_{\langle \alpha, \beta \rangle}(0) = \alpha$ and $f_{\langle \alpha, \beta \rangle}(1) = \beta$. The map sending $\langle \alpha, \beta \rangle$ to $f_{\langle \alpha, \beta \rangle}$ is a bijection between these two forms of κ^2.

Third, we can view κ^2 as the result of cardinal exponentiation. That is, κ^2 (the cardinal operation) is equal to $|\kappa^2|$ (the cardinality of the set of functions from 2 into κ) which we know is the same as $|\kappa^2|$ (the cardinality of the set of ordered pairs) which we know is equal to κ when κ is infinite. Therefore, for infinite κ, we have $\kappa^2 = \kappa$.

One can make similar comments about the interpretations of κ^3, κ^4 and so on. As long as κ is infinite, we have $\kappa^n = \kappa$ for each $n \in \omega$ with $n \geq 1$. For the exponent 0, we have $\kappa^0 = 1$ because κ^0 gives the cardinality of the set of maps from 0 to κ. But, there is only one map with domain 0 (i.e. with an empty domain) and that is the empty map.

At this point, it may seems that we will be able to calculate all of the cardinal exponential values in a similar manner to the way we calculated all the cardinal addition and multiplication facts. Unfortunately, we will show that this impression is incorrect. However, we can make a bit more progress on calculating, or at at least collapsing, many of the remaining cases of cardinal exponentiation.

Lemma 5.18. *For any sets A and B, $|A^B| = |A|^{|B|}$.*

Proof. To show this equality, we need to define a bijection between the set of all functions mapping $B \to A$ and the set of all functions mapping $|B| \to |A|$. Fix bijections $g : B \to |B|$ and $h : A \to |A|$. Because these functions are bijections, there inverses $g^{-1} : |B| \to B$ and $h^{-1} : |A| \to A$ are bijections as well.

We define a function $\Delta : A^B \to |A|^{|B|}$ as follows. Notice that an input to Δ is a function of the form $f : B \to A$ and an output $\Delta(f)$ needs to be a function $\Delta(f) : |B| \to |A|$. For a given function $f : B \to A$, define $\Delta(f)$ by

$$\Delta(f) : |B| \to |A| \text{ with } \Delta(f) = h \circ f \circ g^{-1}.$$

Since g^{-1} maps $|B| \to B$, then f maps $B \to A$ and then h maps $A \to |A|$, this composition gives a function with the correct domain and range. Thus, $\Delta : A^B \to |A|^{|B|}$.

Rather than prove directly that Δ is a bijection, we explicitly define the inverse Δ^{-1} of Δ. That is, $\Delta^{-1} : |A|^{|B|} \to B^A$ needs to satisfy $\Delta^{-1} \circ \Delta : A^B \to A^B$ is the identity function on A^B. An input to Δ^{-1} has the form $\widehat{f} : |B| \to |A|$ and the corresponding output $\Delta^{-1}(\widehat{f})$ has to be a function $B \to A$. Given a function $\widehat{f} : |B| \to |A|$, we define $\Delta^{-1}(\widehat{f})$ by

$$\Delta^{-1}(\widehat{f}) : B \to A \text{ with } \Delta^{-1}(\widehat{f}) = h^{-1} \circ \widehat{f} \circ g.$$

Since g maps $B \to |B|$, then $\widehat{f}$ maps $|B| \to |A|$ and then h^{-1} maps $|A| \to A$, the composition gives function with the correct domain and range. Thus, $\Delta^{-1} : |A|^{|B|} \to A^B$.

It remains to check that $\Delta^{-1} \circ \Delta : A^B \to A^B$ is the identity function. Fix $f : A \to B$ and we calculate as follows.

$$(\Delta^{-1} \circ \Delta)(f) = \Delta^{-1}(\Delta(f)) = \Delta^{-1}(h \circ f \circ g^{-1}) = h^{-1} \circ h \circ f \circ g^{-1} \circ g$$

But, $h^{-1} \circ h : A \to A$ is the identity function Id_A on A and $g^{-1} \circ g : B \to B$ is the identity function Id_B on B. Therefore, this composition reduces to $\mathrm{Id}_A \circ f \circ \mathrm{Id}_B$ which is just f. □

Remember that we have reduced calculating cardinal exponentiation to the case when $2 \leq \kappa, \lambda$ and one of these cardinals is infinite. The next lemma collapses all of these cases when the exponent at least as large as the base.

Lemma 5.19. *Let κ and λ be cardinals. If λ is infinite and $2 \leq \kappa \leq \lambda$, then $\kappa^\lambda = 2^\lambda$.*

Proof. Because $2 \leq \kappa \leq \lambda$, we have $2 \subseteq \kappa \subseteq \lambda$ and therefore $2^\lambda \subseteq \kappa^\lambda \subseteq \lambda^\lambda$ as sets of functions. Hence we have $|2^\lambda| \leq |\kappa^\lambda| \leq |\lambda^\lambda|$ or in other words $2^\lambda \leq \kappa^\lambda \leq \lambda^\lambda$ as cardinal exponentiation.

An element of λ^λ (as a set of functions) is a function $f : \lambda \to \lambda$. We can equate this function f with its graph, namely the set $X_f = \{\langle \alpha, f(\alpha) \rangle \mid \alpha < \lambda\}$. Because f maps into λ, i.e. $f(\alpha) \in \lambda$, we have that $X_f \subseteq \lambda \times \lambda$. Therefore, each function $f \in \lambda^\lambda$ can be viewed as a subset of $\lambda \times \lambda$. In other words, $\lambda^\lambda \subseteq \mathcal{P}(\lambda \times \lambda)$ and hence $|\lambda^\lambda| \leq |\mathcal{P}(\lambda \times \lambda)|$.

Putting these pieces together, we have

$$2^\lambda \leq \kappa^\lambda \leq \lambda^\lambda \leq |\mathcal{P}(\lambda \times \lambda)| = |2^{\lambda \times \lambda}| = |2|^{|\lambda \times \lambda|} = 2^{\lambda \otimes \lambda} = 2^\lambda$$

Therefore, $2^\lambda \leq \kappa^\lambda \leq 2^\lambda$ and hence $2^\lambda = \kappa^\lambda$. □

The next technical lemma is proved by function manipulations similar to those used in Lemma 5.18 and I will leave you to think about the proof.

Lemma 5.20. *Let κ, λ and μ be cardinals. Then*

$$\kappa^{\lambda \oplus \mu} = \kappa^{\lambda} \otimes \kappa^{\mu} \text{ and } \left(\kappa^{\lambda}\right)^{\mu} = \kappa^{\lambda \otimes \mu}$$

Before continuing to discuss our calculation of cardinal exponentiation, we give one (perhaps surprising) application of the results so far.

Theorem 5.21. *Let $\mathcal{C}(\mathbb{R}) = \{f : \mathbb{R} \to \mathbb{R} \mid f \text{ is continuous}\}$. $\mathcal{C}(\mathbb{R})$ has size continuum, that is, $|\mathcal{C}(\mathbb{R})| = |\mathbb{R}| = 2^{\aleph_0}$.*

Proof. Define $\Delta : \mathcal{C}(\mathbb{R}) \to \mathbb{R}^{\mathbb{Q}}$ by $\Delta(f) = f \restriction \mathbb{Q}$. In other words, Δ takes a continuous function on the reals as an input and then outputs the same function but with the domain restricted to the rationals. Note that if f and g are continuous functions on the reals which take the same values on the rationals, i.e. $f \restriction \mathbb{Q} = g \restriction \mathbb{Q}$, then $f = g$. Thus, Δ is one-to-one which means that $|\mathcal{C}(\mathbb{R})| \leq |\mathbb{R}^{\mathbb{Q}}|$. We can now bound the size of $|\mathcal{C}(\mathbb{R})|$ by

$$|\mathcal{C}(\mathbb{R})| \leq |\mathbb{R}^{\mathbb{Q}}| = |\mathbb{R}|^{|\mathbb{Q}|} = \left(2^{\aleph_0}\right)^{\aleph_0} = 2^{\aleph_0 \otimes \aleph_0} = 2^{\aleph_0}.$$

The first inequality was explained above. The first equality follows from Lemma 5.18, the second equality follows from $|\mathbb{R}| = 2^{\aleph_0}$, the third equality follows from Lemma 5.20 and the last equality follows from Theorem 5.13.

On the other hand, for each $r \in \mathbb{R}$, the constant function $f_r(x) = r$ is continuous. Therefore, $|\mathbb{R}| \leq |\mathcal{C}(\mathbb{R})|$ and hence $2^{\aleph_0} \leq |\mathcal{C}(\mathbb{R})|$. Having shown both inequalities, we can conclude that $2^{\aleph_0} = |\mathcal{C}(\mathbb{R})|$. □

To finish our discussion of cardinal exponentiation, we need one more set theoretic tool.

Definition 5.22. Let $\alpha, \beta \in \mathbb{ON}$. A function $f : \beta \to \alpha$ is *cofinal* if range(f) is unbounded in α. That is, for every $\gamma < \alpha$, there is a $\delta < \beta$ such that $\gamma \leq f(\delta)$.

Example 5.23. The identity function $\mathrm{Id}_\alpha : \alpha \to \alpha$ is always cofinal. For some nontrivial examples, the functions $f : \omega \to \omega + \omega$ given by $f(n) = \omega + n$ and $g : \omega \to \omega \cdot \omega$ given by $g(n) = \omega \cdot n$ are each cofinal.

We will be most interested in cofinal maps of the form $f : \alpha \to \kappa$ where κ is a cardinal. Our next examples illustrate two different types of behavior.

Example 5.24. The function $f : \omega \to \aleph_\omega$ defined by $f(n) = \aleph_n$ is cofinal. To see this, fix $\gamma < \aleph_\omega$. We need to find $\delta \in \omega$ such that $\gamma \leq f(\delta)$. Recall that $\aleph_\omega = \sup(\{\aleph_n \mid n \in \omega\})$ and hence $\aleph_\omega$ is the least ordinal greater than each $\aleph_n$. Since $\gamma < \aleph_\omega$, there must be some $n \in \omega$ such that $\gamma \leq \aleph_n$. We fix such an n and set $\delta = n$. Then $f(n) = \aleph_n$ and we have $\gamma \leq \aleph_n = f(n)$.

Example 5.25. There is no cofinal map $f : n \to \omega$ for any $n < \omega$. To see why, fix $n < \omega$ and $f : n \to \omega$. If $n = 0$, then f is the empty function which can never be cofinal. So, assume $n > 0$. Since n is finite, we have range$(f) \subseteq \omega$ is finite and therefore has a greatest element. Since ω does not have a greatest element, range(f) is not unbounded in ω and hence f is not cofinal.

Example 5.26. There is no cofinal map $f : \alpha \to \aleph_1$ for any countable ordinal α. To see why, suppose for a contradiction that there is a countable α and a cofinal map $f : \alpha \to \aleph_1$. We claim that

$$\aleph_1 = \bigcup_{\beta<\alpha} f(\beta)$$

We prove this claim by showing each side of this equation is a subset of the other side. For each $\beta < \alpha$, since $f(\beta) \in \aleph_1$, we have $f(\beta) \subseteq \aleph_1$. Therefore, $\bigcup_{\beta<\alpha} f(\beta) \subseteq \aleph_1$. On the other hand, if $\gamma < \aleph_1$, then $\gamma + 1 \in \aleph_1$ because $\aleph_1$ is a limit ordinal. By the cofinality of f, there is a $\beta < \alpha$ such that $\gamma + 1 \leq f(\beta)$ and hence $\gamma < f(\beta)$. Therefore, for every $\gamma < \aleph_1$, there is a $\beta < \alpha$ such that $\gamma \in f(\beta)$. It follows that $\aleph_1 \subseteq \bigcup_{\beta<\alpha} f(\beta)$. This proves the claim.

For each $\beta < \alpha$, $f(\beta) \in \aleph_1$ and hence $f(\beta)$ is countable because $\aleph_1$ is the least uncountable ordinal. Since α is countable, the union $\bigcup_{\beta<\alpha} f(\beta)$ is a countable union of countable sets. Therefore, the displayed equation proved in the claim says that $\aleph_1$ is a countable union of countable sets. But, we know a countable union of countable sets is countable, whereas $\aleph_1$ is uncountable. This gives the desired contradiction.

Definition 5.27. Let $\alpha \in \mathbb{ON}$. The *cofinality of* α, denoted $\mathrm{cf}(\alpha)$, is the least ordinal β such that there is a cofinal map $f : \beta \to \alpha$.

We will typically be concerned with calculating cofinalities of cardinals. Because the identity function $\mathrm{Id}_\kappa : \kappa \to \kappa$ is cofinal, we have $\mathrm{cf}(\kappa) \leq \kappa$. By Example 5.25, $\mathrm{cf}(\aleph_0) = \aleph_0$ and by Example 5.26, $\mathrm{cf}(\aleph_1) = \aleph_1$. However, by Example 5.24, $\mathrm{cf}(\aleph_\omega) = \omega$. (The example shows $\mathrm{cf}(\aleph_\omega) \leq \omega$. But, $\mathrm{cf}(\aleph_\omega)$ cannot be finite because $\aleph_\omega$ does not have a greatest element.)

Definition 5.28. An infinit cardinal κ is called *regular* if $\mathrm{cf}(\kappa) = \kappa$. Otherwise, κ is called *singular*.

Example 5.29. $\aleph_0$ and $\aleph_1$ are regular cardinals while $\aleph_\omega$ is a singular cardinal.

Theorem 5.30 (König's Theorem)**.** *For every infinite cardinal κ, we have $\kappa^{cf(\kappa)} > \kappa$.*

Proof. Since $\mathrm{cf}(\kappa) \geq \omega$ because κ is a limit ordinal, we have $\kappa^{\mathrm{cf}(\kappa)} \geq \kappa$. To show this inequality is strict, we need to show that there is no bijection from κ to $\kappa^{\mathrm{cf}(\kappa)}$. To do this, we show that every one-to-one map from κ into $\kappa^{\mathrm{cf}(\kappa)}$ is not onto.

Fix an arbitrary one-to-one function $G : \kappa \to \kappa^{\mathrm{cf}(\kappa)}$. The function G takes an ordinal $\beta < \kappa$ as an input and the output $G(\beta)$ is a function from $\mathrm{cf}(\kappa)$ to κ. That is, $G(\beta) : \mathrm{cf}(\kappa) \to \kappa$, and so we can apply $G(\beta)$ to any ordinal $\alpha < \mathrm{cf}(\kappa)$ to get $G(\beta)(\alpha) \in \kappa$. At the cost of possible overemphasis, note that $G(\beta)(\alpha)$ is formed by plugging $\beta < \kappa$ into G to obtain the function $G(\beta) : \mathrm{cf}(\kappa) \to \kappa$ and then plugging $\alpha < \mathrm{cf}(\kappa)$ into this function $G(\beta)$ to obtain $G(\beta)(\alpha) \in \kappa$.

To show that G is not onto, we need to give an element of $\kappa^{\mathrm{cf}(\kappa)}$ which is not in the range of G. That is, we need to specify a function $h : \mathrm{cf}(\kappa) \to \kappa$ such that h is not in the range of G. Unraveling one step further, to show that h is not in the range of G means that for every $\beta < \kappa$, we have $h \neq G(\beta)$ as functions. In other words, there is an $\alpha < \mathrm{cf}(\kappa)$ such that $h(\alpha) \neq G(\beta)(\alpha)$. Summarizing, having fixed $G : \kappa \to \kappa^{\mathrm{cf}(\kappa)}$, we need to define a function $h : \mathrm{cf}(\kappa) \to \kappa$ such that for every $\beta < \kappa$, there is an $\alpha < \mathrm{cf}(\kappa)$ such that $h(\alpha) \neq G(\beta)(\alpha)$.

To define h, we fix a cofinal map $f : \mathrm{cf}(\kappa) \to \kappa$. Define $h : \mathrm{cf}(\kappa) \to \kappa$ by

$$h(\alpha) = \min(\kappa \setminus \{G(\beta)(\alpha) \mid \beta < f(\alpha)\}).$$

We need to check that this definition makes sense. Because κ is a cardinal, and hence an ordinal, it is well ordered. So, there is a minimal element of $\kappa \setminus \{G(\beta)(\alpha) \mid \beta < f(\alpha)\}$ as long as this set is not empty. To see that this set is nonempty, it suffices to show that $|\{G(\beta)(\alpha) \mid \beta < f(\alpha)\}| < \kappa$.

To see this inequality, consider the map $g : f(\alpha) \to \{G(\beta)(\alpha) \mid \beta < f(\alpha)\}$ given by $g(\beta) = G(\beta)(\alpha)$. By definition, g is onto and hence $|\{G(\beta)(\alpha) \mid \beta < f(\alpha)\}| \leq |f(\alpha)|$. But, $f(\alpha) \in \kappa$ and κ is a cardinal, so $|f(\alpha)| < \kappa$. Therefore, $|\{G(\beta)(\alpha) \mid \beta < f(\alpha)\}| < \kappa$ as required to show that the function h is defined.

Finally, we need to show that h is not in the range of G. Suppose for a contradiction that h is in the range of G. Then $h = G(\beta)$ for some $\beta < \kappa$. This means that $h(\alpha) = G(\beta)(\alpha)$ for all $\alpha < \mathrm{cf}(\kappa)$. To arrive at a contradiction, notice that since $f : \mathrm{cf}(\kappa) \to \kappa$ is cofinal, there is some $\alpha < \mathrm{cf}(\kappa)$ such that $\beta < f(\alpha)$. Consider the value of $G(\beta)(\alpha)$. By definition, $h(\alpha) \neq G(\gamma)(\alpha)$ for all $\gamma < f(\alpha)$. But, $\beta < f(\alpha)$ and hence $h(\alpha) \neq G(\beta)(\alpha)$. This contradicts the fact noted above that $h(\alpha) = G(\beta)(\alpha)$ for all $\alpha < \mathrm{cf}(\kappa)$. □

Corollary 5.31. *The cofinality of* $2^{\aleph_0}$ *satisfies* $\mathrm{cf}(2^{\aleph_0}) > \aleph_0$.

Proof. Assume for a contradiction that $\mathrm{cf}(2^{\aleph_0}) = \aleph_0$. Consider the following calculation.

$$\left(2^{\aleph_0}\right)^{\mathrm{cf}(2^{\aleph_0})} = \left(2^{\aleph_0}\right)^{\aleph_0} = 2^{\aleph_0 \otimes \aleph_0} = 2^{\aleph_0}$$

The first equality follows from our assumption, the second equality follows from Lemma 5.20 and the third equality follows from Theorem 5.13. By this calculation,

$$\left(2^{\aleph_0}\right)^{\mathrm{cf}(2^{\aleph_0})} = 2^{\aleph_0}$$

which directly contradicts König's Theorem. □

Finally, let us return to what we can say about the values of cardinal exponentiation κ^λ. The simplest case we have not determined is $2^{\aleph_0}$. By Cantor's Theorem, we know $2^{\aleph_0} > \aleph_0$ and König's Theorem, we know $\mathrm{cf}(2^{\aleph_0}) > \aleph_0$. Unfortunately, this is essentially all that we can say about the value of $2^{\aleph_0}$. For any cardinal κ such that $\kappa > \aleph_0$ and $\mathrm{cf}(\kappa) > \aleph_0$, one can use Cohen's method of forcing to construct a model of ZFC in which $2^{\aleph_0} = \kappa$. Giving such a forcing construction is beyond the scope of this course, but it is worth mentioning that you now understand all the constraints on the value of $2^{\aleph_0}$ within our best mathematical model of set theory.

More generally, using Cantor's Theorem, we can define the following hierarchy of infinite cardinals indexed by the ordinals.

$$\beth_0 = \omega$$
$$\beth_{\alpha+1} = 2^{\beth_\alpha}$$
$$\beth_\alpha = \sup(\{\beth_\beta \mid \beta < \alpha\}) \text{ for limit } \alpha$$

We know that $\aleph_0 = \beth_0$, but after that we are left in limbo. The statement that $\aleph_1 = \beth_1$, i.e. that $2^{\aleph_0} = \aleph_1$ or equivalently that $|\mathbb{R}| = \aleph_1$, is called the *continuum hypothesis* and is abbreviated CH. Gödel proved that CH was consistent with ZFC and Cohen proved that ¬CH was consistent with ZFC.

The more general statement that $\beth_\alpha = \aleph_\alpha$ for all ordinals α is called the *generalized continuum hypothesis* and abbreviated GCH. Gödel proved that GCH is consistent with ZFC, but obviously since ¬CH is consistent with ZFC, so is ¬GCH.