What exactly does "infinite" mean? We begin our investigation by first constructing the set of natural numbers, which we denote by ω (as opposed to the more typical $\mathbb{N}$ since we want to make explicit the construction we have in mind), and use this to study the properties of finite and infinite sets. After this, we look at the "smallest" infinite sets, and prove that there are larger ones. From here, we move to the discussion of ordinals and cardinals, which provide canonical representatives of order and size, respectively. We end by looking at some of the properties of cardinal arithmetic.

Section 1

The Natural Numbers

To talk about finite sets requires us to talk about the Natural Numbers, so we start by defining them and discussing their elementary properties in terms of this construction.

We can 'construct' the set of natural numbers using the axioms of ZFC . Our idea for doing this is based on defining, recursively, the natural numbers. We begin with $0 = \varnothing$, and given a natural number n, we define $n + 1$ by $n + 1 := n^+ := n \cup \{n\}$, called the **successor** of n.

There are plenty of sets that contain the natural numbers, but ω is the *smallest* set containing all of them. Making this precise, we say that a set I is **inductive** if $\varnothing \in I$ and for every $x \in I$, x^+ is also in I. We wish to define ω to be the intersection of all these inductive sets:

$$\omega := \bigcap X$$

where X is the set of all inductive sets. However, from a technical standpoint the "set" X actually isn't a set at all, giving us some issues. Moreover, we don't know whether there actually are any inductive sets. We solve both of these problems by appealing to the **Axiom of Infinity**, which states that an inductive set I exists. Then we can consider X to be the set of all inductive subsets of I.

This definition is incredibly useful, because it allows us to capture one of the most important properties of ω: the **Principle of Mathematical Induction**, which states that if a property P is true for 0, and if being true for n implies that it is true for $n + 1$, then P is true for *every* natural number.

In terms of inductive sets, if we consider the set $S = \{n \in \omega \mid n \text{ has the property } P\}$, if we can show that S is inductive, then we can conclude that $S = \omega$. This will be our mindset when using induction.

Section 2

Finite Sets

The existence of ω also allows us to define precisely what it means for a set to be finite: we say that a set S is **finite** if there exists natural number $n \in \omega$ such that there exists a bijection from S onto n; n is called the **cardinality** or **size** of S.

We shall examine some of the properties of finite sets, including that the cardinality of a finite set is well-defined, that subsets of finite sets are finite, that unions and cartesian products of two finite sets are finite, and finally that the power set of a finite set is finite. Our main methodology will be by building up a subset consisting of all natural numbers for which the relevant property holds (e.g. that any subset of a set of size n is finite). We then show that this set is inductive, from which we find that it is equal to ω.

However, in order to actually define these sets, we shall be making extensive use of the Axiom Schema of Comprehension. For this reason, we need to build up some background in which properties of sets can be expressed by a wff.

Now we proceed to examine some of the properties of finite sets, starting with proof that the cardinality of a finite set is well-defined:

Theorem 2.1

Suppose that there exists a bijection from S to n and a bijection from S to m, where $n, m \in \omega$. Then $n = m$.

Proof.
It suffices to show that there exists no bijection between n and m for $n \neq m$, as if f is the bijection $f : S \to n$ and g is the bijection $g : S \to m$, then $g \circ f^{-1}$ gives us a bijection from n to m, from which we draw the conclusion that $n = m$.

We let

$$S = \{n \in \omega \mid \text{there exists no bijection from } n \text{ to } m \text{ for } m \in \omega \text{ and } m \neq n\}.$$

We wish to show that S is inductive, from which it will follow that $S = \omega$. To do this, we begin by showing that there exists no non-empty set X such that there is a bijection f between X and $\varnothing$, which follows because since X is non-empty there exists $x \in X$, and there does not exist any $y \in \varnothing$ such that $f(y) = x$ (because there are no elements of $\varnothing$ at all). Thus, $\varnothing \in S$.

Now suppose that $n \in S$; we wish to show that $n^+ \in S$. Suppose for the sake of a contradiction that there exists a bijection between n^+ and a natural number distinct from n^+; we can assume that this natural number is a successor of another by our earlier comments, so suppose $f : n^+ \to m^+$ is a bijection where $n \neq m$. That is, we have a bijection from $n \cup \{n\}$ to $m \cup \{m\}$. If $f(n) = m$, then we can restrict f to n to give us a bijection from n to m, giving us a contradiction, since $n^+ \neq m^+$ implies $n \neq m$. Otherwise, there is ℓ in m such that $f(n) = \ell$. We define a map $g : m^+ \to m^+$ that is the identity on $m^+ \smallsetminus \{\ell, m\}$, maps m to ℓ, and maps ℓ to m. This is a bijection as it is the disjoint union of two bijections. Then $h = g \circ f$ gives us a bijection $h : n \cup \{n\} \to m \cup \{m\}$ that sends n to m, so that we may restrict to n to give us a bijection from n to m, giving us the necessary contradiction. Thus, we find that $n^+ \in S$.

Hence, S is inductive, and thus $\omega = S$. □

Natural things we would like our finiteness to have includes the fact that subsets of finite sets are finite and have size less than that of the superset, that the union of two finite sets is finite, the cartesian product of two finite sets is finite, and that the power set of a finite set is finite:

Proposition 2.2

Let x be a finite set, and y a subset of x. Then y is finite.

Proof.
Let

$$S = \{n \in \omega \mid \text{every subset of a set of size } n \text{ is finite}\}.$$

Clearly $\varnothing$ is in S because the only subset of $\varnothing$ is $\varnothing$ itself, which is finite.

Now suppose that $n \in S$, and let x be a set of size n^+ and y a subset of x. Since x has size n^+, there is a bijection $f : x \to n^+$, so let z be the element of x mapped to n. If y does not contain z, then $y \subset x \smallsetminus \{z\}$, which has size n (the bijection is given by the restriction $f|_{x \smallsetminus \{z\}}$), so that because $n \in S$ by hypothesis, we find that y is finite. On the other hand, if z is in y, then $y \smallsetminus \{z\}$ is finite by the above comments, and thus there is a natural number m and a bijection $g : y \smallsetminus \{z\} \to m$. Then we can extend g to $g' : y \to m^+$ defined by setting $g'(w) = g(w)$ when $w \in y \smallsetminus \{z\}$ and $g'(z) = m$, showing us that y is finite.

Thus, S is inductive and $S = \omega$. □

Proposition 2.3

If x is a finite set and y is a proper subset, then the size of y is strictly less than that of x.

Proof.
Let

$$S = \{n \in \omega \mid \text{all proper subsets of a set of size } n \text{ have size less than } n\}.$$

$\varnothing \in S$ trivially because $\varnothing$ has no proper subsets. Now suppose $n \in \omega$, and let x have size n^+, and $f : x \to n^+$ a bijection. Let y be any proper subset of x, and let z be an element of x not in y. Let $w \in x$ be the unique element such that $f(w) = n$, and let $g : x \to x$ be the bijection that fixes all elements of $x \smallsetminus \{w, z\}$, sends w to z, and z to w. Then $g \circ f : x \to n^+$ is a bijection that sends z to n, and by restricting g we find that $x \smallsetminus \{z\}$ has size n. Now, y is a subset of $x \smallsetminus \{z\}$.

Case 1: If $y = x \smallsetminus \{z\}$, then y has size n which is less than n^+.

Case 2: If $y \neq x \smallsetminus \{z\}$, then y is a proper subset of a set of size n, and thus we find that y has size less than n, and thus has size less than n^+.

Thus $n^+ \in S$, S is inductive, and $S = \omega$. □

PROPOSITION 2.4

Suppose that x and y are finite sets. Then $x \cup y$ is also finite.

Proof.
Let

$$S = \{n \in \omega \mid \text{the union of a set of size } n \text{ and a finite set is always finite}\}.$$

It is clear that $\varnothing \in S$ because $A \cup \varnothing = A$, so that if A is finite, then $A \cup \varnothing$ is also finite.

Now suppose that $n \in S$, that x has size n^+, and y is finite. Because x has size n^+, there exists a bijection $f : x \to n^+$, and let $z \in x$ be the element mapped to n. Then the restriction $f|_{x \smallsetminus \{z\}} : x \to n$ is also a bijection. Now, since $n \in S$ we know that $(x \smallsetminus \{z\}) \cup y$ is finite, and thus there exists m and a bijection $g : (x \smallsetminus \{z\}) \cup y \to m$. If $z \in y$, then $(x \smallsetminus \{z\}) \cup y = x \cup y$, and thus $x \cup y$ is finite. On the other hand, if $z \notin y$, then we can extend g to $g' : x \cup y \to m^+$ defined by $w \mapsto g(w)$ if $w \in (x \smallsetminus \{z\}) \cup y$ and $z \mapsto m$. Thus, we find that $x \cup y$ is finite in either case, and $n^+ \in S$.

Hence S is inductive and $S = \omega$. □

PROPOSITION 2.5

Suppose that x and y are finite sets. Then $x \times y$ is also finite.

Proof.
Let

$$S = \{n \in \omega \mid \text{the product of a set of size } n \text{ and a finite set is finite}\}.$$

$\varnothing \in S$ because $\varnothing \times A = \varnothing$ for every set A.

Now suppose that $n \in S$, let x be a set of size n^+, and y a finite set. There exists a bijection $f : x \to n^+$, and letting z be the unique element in x such that $f(z) = n$, the restriction $f|_{x \smallsetminus \{z\}} : x \smallsetminus \{z\} \to n$ is a bijection. Then $(x \smallsetminus \{z\}) \times y = (x \times y) \smallsetminus (\{z\} \times y)$ is finite by the hypothesis that $n \in S$. $\{z\} \times y \cong y$, so that $\{z\} \times y$ is finite. Then $x \times y = [(x \smallsetminus \{z\}) \times y] \cup [\{z\} \times y]$ is the union of two finite sets, and thus is finite, showing us that $n^+ \in S$.

This shows us that S is inductive, and thus $S = \omega$. □

PROPOSITION 2.6

Suppose that x is finite. Then the power set $\mathcal{P}(x)$ is finite.

Proof.
Let

$$S = \{n \in \omega \mid \text{the power set of a set of size } n \text{ is finite}\}.$$

$\varnothing \in S$ because $\mathcal{P}(\varnothing) = \varnothing$ which is finite.

Now suppose that $n \in S$, and let x be a set of size n^+. There exists a bijection $f : x \to n^+$, and letting z be the unique element of x such that $f(z) = n$, the restriction $f|_{x \setminus \{z\}} : x \setminus \{z\} \to n$ is a bijection. Then $x \setminus \{z\}$ is of size n, and thus $\mathcal{P}(x \setminus \{z\})$ is finite.

We can write $\mathcal{P}(x)$ as

$$\mathcal{P}(x) = \mathcal{P}(x \setminus \{z\}) \cup \{w \cup \{z\} \mid w \in \mathcal{P}(x \setminus \{z\})\}$$

because a given subset v of x can either contain z, in which case v can be written $\{z\} \cup (v \setminus \{z\})$ with $v \setminus \{z\}$ a subset of $x \setminus \{z\}$, or it does not contain z, in which case it is a subset of $x \setminus \{z\}$. Note that the collection $\{w \cup \{z\} \mid w \in \mathcal{P}(x \setminus \{z\})\}$ is a set by the Axiom of Replacement.

Next, we show that $\{w \cup \{z\} \mid w \in \mathcal{P}(x \setminus \{z\})\}$ is in bijection with $\mathcal{P}(x \setminus \{z\})$ by defining the bijection $g : \mathcal{P}(x \setminus \{z\}) \to \{w \cup \{z\} \mid w \in \mathcal{P}(x \setminus \{z\})\}$ by $g : w \mapsto w \cup \{z\}$. This is an injection because if $w \cup \{z\} = w' \cup \{z\}$, then because $z \notin w, w'$ we find that $w = w'$; that it is a surjection follows immediately from the definition of the codomain; thus it is a bijection, and because the domain is finite, the codomain is finite as well. Then $\mathcal{P}(x)$ is the union of two finite sets, and is thus finite itself. □

Section 3

Infinite Sets

Now we show that there exist infinite sets, namely sets that are not in bijection with a natural number. The obvious choice for such a set is ω itself.

We begin with the following theorem, which helps us characterize finite sets:

Theorem 3.1

Suppose that the finite sets x and y have equal size, and let $f : x \to y$ be any function. Then the following are equivalent:

(i) f is injective.

(ii) f is surjective.

(iii) f is bijective.

Proof.
(i) $\implies$ *(ii)* Let x have size n and y have size m, and let $g : x \to n$ and $h : y \to m$. By restricting $f : x \to y$ to $f[x]$, we find that $f[x]$ has the same size as x. But since x has the same size as y, it follows that $f[x]$ has the same size as y. This implies that $f[x] = y$, as if either $f[x] = y$ or it is a proper subset, but we know that a proper subset of a finite set is also finite with size less than that of the superset. This is precisely that f is surjective.
(ii) $\implies$ *(iii)* Let $f : x \to y$ be surjective, and let $<$ be a strict well-order on x. We define a function $g : y \to x$ by defining $g(w)$ to be the least element of $f^{-1}(w)$, which exists because x is well-ordered and $f^{-1}(w)$ is non-empty for every w because f is surjective. In particular, $g \circ f = \mathrm{id}_x$.

g is an injection because if $g(z) = g(z') = w$, then by definition $f(z) = f(z') = w$. But z and z' are the minimum such values in x, so that they must be equal by the uniqueness of the minimum. We have already shown that because x and y have equal size and g is an injection, it is also a surjection, and hence a bijection. But then because $g \circ f = \mathrm{id}_x$, we find that $f = g^{-1}$ and f is a bijection.
(iii) $\implies$ *(i)* This follows immediately from the fact that bijectivity means injectivity and surjectivity.

□

We say that a set x is **Dedekind finite** if given any function $f : x \to x$, injectivity, surjectivity, and bijectivity are all equivalent. Then the above theorem shows us that if x is finite, then x is Dedekind finite.

It ends up that being Dedekind finite implies being finite, though we do not prove this now.

In particular, it provides us with enough information to show that ω is infinite:

Corollary 3.2

ω is infinite.

Proof.
We define a map $f : \omega \to \omega$ by $f : n \mapsto n^+$. This is injective, as if $n^+ = m^+$, then $n = m$. However, it is not surjective, because there does not exist any n such that $n^+ = \varnothing$.

Thus, if ω was finite, it would be Dedekind finite and thus injectivity would imply surjectivity, giving us a contradiction. □

Just as we were able to show that the union of two finite sets was finite, we may show that any union in which at least one of the sets being unioned is infinite is also infinite. Moreover, we can show that if a family of sets is infinite, then its union is also infinite.

Proposition 3.3

(a) Let x be infinite and $x \subset y$. Then y is infinite.

(b) Let X be a family of sets in which at least one of the elements of X is infinite. Then $\bigcup_{x \in X} x$ is infinite.

(c) Let X be an infinite family of sets. Then $\bigcup_{x \in X} x$ is infinite.

Proof.

(a) If y is finite, then it follows that x is finite, giving us a contradiction.

(b) Let y denote the infinite set in X. Then $y \subset \bigcup_{x \in X} x$, so by (a) we find that $\bigcup_{x \in X} x$ is infinite.

(c) Suppose that the union $U = \bigcup_{x \in X} x$ is finite. Then $\mathcal{P}(U)$ is finite by **??**; we must show that $X \subset \mathcal{P}(U)$, from which point X will be finite, giving us a contradiction. Let y be any element of X. Because $y \subset U$, it follows that $y \in \mathcal{P}(U)$, and thus $X \subset \mathcal{P}(U)$ and we have our contradiction. Thus, U must be infinite.

□

Corollary 3.4

Every inductive set is infinite.

Proof.
Every inductive set contains ω, so that since ω is infinite, every inductive set is also infinite. □

It is for this reason that the Axiom of Infinity has its name: it asserts the existence of an infinite set.

Section 4

Cardinality and Countability

With the introduction of the natural numbers, we were able to associate with every finite set a unique 'size' or 'cardinality'. With the introduction to infinite sets, we encountered sets for which there did not exist a natural number that could be associated via the existence of a bijection. In this section we look at some of the properties of the concept of cardinality as a measure of size, Transfinite Recursion, and one particular class of infinite sets in bijection with one-another, the countably-infinite sets.

Subsection 4.1

Cardinality and the Cantor-Schroeder-Bernstein Theorem

We extend our use of 'cardinality' to all sets by saying that two sets x and y *have the same cardinality* if there exists a bijection $f : x \to y$, which we denote as $|x| = |y|$. We define an order $\leq$ by saying that $|x| \leq |y|$ if there exists an injection $f : x \to y$, and write $|x| < |y|$ if there exists such an injection but no bijection. These orders represent a comparison of the sizes of x and y; if there is an injection of x into y, then we are in a sense saying that we can fit a copy of x into y (because x is in bijection with a subset of y by restricting the injection to its image), while the surjectivity says that it x is big enough to cover y entirely.

This ordering has the nice property of being antisymmetric. First, we prove a lemma (which is an important result in and of itself):

Lemma 4.1: Knaster-Tarski Fixed Point Theorem

Let P be a poset in which every subset has a least upper bound, and let $f : P \to P$ be a monotonic function. Then f has a fixed point, i.e. a point $x \in P$ such that $f(x) = x$.

Proof.
The main idea of the proof is that if $f(y) = y$, then $y \leq f(y)$ and $f(y) \leq y$. Since f is monotonic, $f(y) \leq f(f(y))$, and since the least upper bound of any subset of P exists, P is bounded above. Thus, our idea is to 'push y up until the inequality $y \leq f(y)$' becomes an equality since f can either keep everything the same or shrink it:

We define
$$S := \{x \in P \mid x \leq f(x)\}.$$

Because f is monotonic, we know that $x \leq f(x)$ implies $f(x) \leq f(f(x))$, so that $f[S] \subset S$. Now let y be the least upper bound of S. By hypothesis, for every $x \in S$ we have $x \leq f(x)$ and $f(x) \in S$. Since f is monotonic, $x \leq y$ implies $f(x) \leq f(y)$, and since $x \leq y$ for every element of x, we find that $x \leq f(x) \leq f(y)$ for every $x \in S$, and thus $f(y)$ is an upper bound of S. Since y is the least upper bound, we have $y \leq f(y)$, and thus $y \in S$. But then y is the maximum of S and thus $f(y) \leq y$ (since S is closed under f), leading us to the conclusion that $f(y) = y$. Thus, y is the required fixed point. □

Theorem 4.2: Cantor-Bernstein-Schroeder Theorem

Suppose there exist injections $f : X \to Y$ and $g : Y \to X$. Then there exists a bijection $h : X \to Y$. That is, if $|X| \leq |Y|$ and $|Y| \leq |X|$, then $|X| = |Y|$.

Proof.
Our methodology will be to piece together the two injections f and g in order to create the required injection by finding pieces of S and T that are 'compatible'. We create a function from $\mathcal{P}(S)$ to itself that is monotonic, so that since $\mathcal{P}(S)$ is a complete lattice, we can use the Knaster-Tarski Fixed Point Theorem:

We define $F : \mathcal{P}(S) \to \mathcal{P}(S)$ by

$$F(A) := S \setminus g[T \setminus f[A]]$$

This is a monotonic function due to the fact that $A \subset B$ implies $f[A] \subset f[B]$ and $T \setminus B \subset T \setminus A$. Then we have

$$A \subset B \Rightarrow f[A] \subset f[B] \Rightarrow T \setminus f[B] \subset T \setminus f[A] \Rightarrow g[T \setminus f[B]] \subset g[T \setminus f[A]] \Rightarrow F(A) \subset F(B).$$

Thus, F is monotonic, and by the Knaster-Tarski Fixed Point Theorem there exists a fixed point $A \subset S$ such that $F(A) = A$. Thus, A is the complement of $g[T \setminus f[A]]$ in S. By restricting f to A, we have a bijection between A and $f[A]$, and by restricting g to $T \setminus f[A]$, we have a bijection between $g[T \setminus f[A]]$ and $T \setminus f[A]$. Since $g[T \setminus f[A]]$ and A are complements, they are disjoint so that we can take $h(x)$ to be defined by

$$h(x) := \left(f|_A \cup g^{-1}|_{g[T \setminus f[A]]}\right)(x) = \begin{cases} f(x) & x \in A \\ g^{-1}(x) & x \in g[T \setminus f[A]] \end{cases}$$

which we know to be a bijection on $A \cup g[T \setminus f[A]] = S$ to $f[A] \cup T \setminus f[A] = T$. □

Subsection 4.2

Transfinite Recursion, Countable Sets, and Uncountable Sets

We say that a set X is **countably infinite** if $|X| = |\omega|$, i.e. if there is a bijection $f : \omega \to X$; a set X is said to be **countable** if it is either countably infinite or finite, or equivalently that there is an injection of X into ω. We shall look at some of the nice facts concerning countability, including that subsets of countable sets countable, that the Cartesian product of countable sets is countable, and that countable unions of countable sets are countable. We follow all of this by showing that there exist sets that are **uncountable**, meaning infinite but not countably-infinite.

We start our discussion by proving a useful fact, both in order to define our necessary bijections and later to axiomatically produce binary operations on sets. This is the Principle of Recursion: given a recursive formula and initial conditions, there exists a function that fulfills the formula and the conditions. To facilitate our proof of this in a general form, we say that a subset T of a well-ordered set S is **downwards closed** if for every $y \in T$ and $x \in S$ such that $x \leq y$ ($x < y$ if X is strict well-ordered), then $x \in T$. Given a well-ordered set S, we denote by $S{\downarrow_x}$ the **initial segment**

$$S{\downarrow_x} := \{y \in S \mid y < x\}.$$

Lemma 4.3

Let T be a downwards closed subset of a well-ordered set S. Then either $S = T$ or $T = S{\downarrow_x}$ for some $x \in X$.

Proof.
Suppose $T \neq S$, and let x be the least element of $S \setminus T \subset S$. Then $S{\downarrow_x} \subset T$ as if $y \in S{\downarrow_x}$ is not in T then we have contradicted our choice of x. If $T \not\subset S{\downarrow_x}$, then there is $x' > x$ such that $x' \in T$, contradicting the fact that T is downwards closed since $x \notin T$ by definition. Thus, $T \subset S{\downarrow_x}$ and thus $T = S{\downarrow_x}$. □

Theorem 4.4: Transfinite Recursion

Let S be a well-ordered set, and suppose we have a function $g : \{(x, h) \mid x \in S, h : S{\downarrow_x} \to T\} \to T$, called a **recursion formula**. Then there exists a unique function $f : S \to T$ such that for every $x \in S$, $f(x) = g(x, f|_{S{\downarrow_x}})$.

Proof.
We say that a function f is *admissible* if its domain is a downwards closed subset of S and it satisfies $f(x) = g(x, f|_{S\downarrow_x})$ for every x in its domain. For example, the function $f : \{s\} \to T$, where s is the least element of S, that sends s to $g(s, \varnothing)$ is admissible because $f|_{S\downarrow_s} = f|_\varnothing = \varnothing$.

Given two admissible functions f_1 and f_2 with domains $S\downarrow_{x_1}$ and $S\downarrow_{x_2}$ and $x \in S\downarrow_{x_1} \cap S\downarrow_{x_2}$, then $f_1(x) = f_2(x)$, as otherwise we could pick the smallest such value s for which $f_1(x) \neq f_2(x)$, so that from the fact that they are admissible we find that $f_1(x) = g(x, f_1|_{S\downarrow_x}) = g(x, f_2|_{S\downarrow_x})$. But by hypothesis, $f_1|_{S\downarrow_x} = f_2|_{S\downarrow_x}$ since x is the least element where they disagree. But then $f_1(x) = f_2(x)$, showing us that f_1 and f_2 agreed at x from the start.

We define f to be the union of the set $\{f' \mid f' \text{ admissible}\}$. We now check that f is an admissible function:

f is a function because given any x in its domain, there is an admissible function f' with a domain that contains x; by the above, any other admissible function defined on x must have the same value, and thus f is well-defined at x.

f has domain equal to the union of downwards closed subsets of S, which is precisely the downwards closed subset $S\downarrow_y$ where y is the least element not in the union, which follows from the fact that if there were x in the union greater than y, then because x lies in some downwards closed subset that makes up the union, y also lies in that downwards closed subset giving us a contradiction. Ify does not exist, then it must be that the domain is S, which is trivially downwards closed. Thus, f has as domain a downwards closed subset.

f satisfies $f(x) = g(x, f|_{S\downarrow_x})$ for each x in its domain due to the fact that f restricts to an admissible function defined on x (as otherwise x would not be in the domain), say f'. Then $f(x) = f'(x) = g(x, f'|_{S\downarrow_x}) = g(x, f|_{S\downarrow_x})$ because the restriction of a restriction $f|_A$ to a set B is simply the restriction $f|_{A\cap B}$.

It only remains to show that f has domain equal to S: otherwise we could pick the smallest value x not in the domain and extend f by defining $f(x) = g(x, f_{S\downarrow_x})$. But this extension is an admissible function, and thus must be in the union that defines f, showing us that f has domain S.

Thus, f is an admissible function with domain S that satisfies $f(x) = g(x, f|_{S\downarrow_x})$ for every $x \in S$. Uniqueness follows from the fact that every other such function must agree with f on all of S, or simply that the two functions are equal. □

The reason that the function $g : \{(x,h) \mid x \in S, h : S\downarrow_x \to T\} \to T$ is called a recursion formula is because the function g tells us what x maps to given we know all the previous values. For example, if we made recursive relation $A_{n^+} = A_n \cup \{A_n\}$, then in order to compute A_{n^+} for a given n, we would need to know the value of A_m for every $m < n^+$; these values are precisely what the function h communicates to us. The function f produced above has the initial condition built in, as $g(s, \varnothing) = f(s)$ for the least element s, and does not depend upon f, but only g.

Proposition 4.5

Let X be a countably-infinite set, and let $Y \subset X$ be infinite. Then Y is countably-infinite.

Proof.
Because X is countable-infinite, there exists a bijection $h : \omega \to X$. With the standard strict well-order on ω, we can create a well-ordering on X by forcing h to be order-preserving, making h into an order isomorphism. We define the following recursion formula:

$$g(n, h : \omega\downarrow_n \to X) \coloneqq \text{least element of } Y \setminus h[\omega\downarrow_n].$$

This is well-defined due to the fact that $h[\omega\downarrow_n]$ is finite and Y is infinite. By Transfinite Recursion, there exists a function $f : \omega \to X$ whose image lies entirely in Y; we show that it is our required bijection.

Suppose $f(n) = f(m)$, and assume without loss of generality that $m \leq n$. If $m < n$, then

$f(n)$ = least element of $Y \setminus f[\omega{\downarrow_n}]$. But $f(m)$ is such a value in $f[\omega{\downarrow_n}]$, giving us a contradiction. Thus, $n = m$, and f is injective.

Let y be the smallest value of Y not in the image of f; then because $\omega{\downarrow_n} = n$ is finite, it follows that $Y{\downarrow_y}$ is finite and lies entirely in the image of f. Because f is injective and $Y{\downarrow_y}$ is finite, $f^{-1}[Y{\downarrow_y}]$ is finite; if $f^{-1}[Y{\downarrow_y}]$ were not finite, then $f|_{f^{-1}[Y{\downarrow_y}]}$ could not be a bijection, but since it is a surjection onto $Y{\downarrow_y}$, it cannot be a injection, giving us a contradiction. Since f is an injection, $f|_{f^{-1}[Y{\downarrow_y}]}[f^{-1}[Y{\downarrow_y}]] = Y{\downarrow_y}$, showing us that $f|_{f^{-1}[Y{\downarrow_y}]}$ is a bijection of $f^{-1}[Y{\downarrow_y}]$ onto $Y{\downarrow_y}$. Since $Y{\downarrow_y}$ is a initial segment and f is an order isomorphism, $f^{-1}[Y{\downarrow_y}]$ is also an initial segment and thus is equal to $\omega{\downarrow_m} = m^+$ for some natural number m (it cannot be ω since it is finite). Then $f(m)$ = least element of $Y \setminus f[\omega{\downarrow_m}] = Y \setminus Y{\downarrow_y}$, which by definition is y. Thus, f is surjective, and f is a bijection. □

Now we move on to the question of finite Cartesian Products of countably-infinite sets, first looking at $\omega \times \omega$.

Theorem 4.6

$\omega \times \omega \cong \omega$.

Proof.
We shall use some simple properties concerning ω by making use of its number-theoretic properties, namely prime factorization. It is clear that $|\omega| \leq |\omega \times \omega|$, so it remains to prove that $|\omega \times \omega| \leq |\omega|$. Define $f : \omega \times \omega \to \omega$ by $f(n, m) = 2^n 3^m$. Then we may apply **??**. □

Corollary 4.7

If X and Y are countably-infinite, then $X \times Y$ is countably-infinite.

Proof.
Because X and Y are countably-infinite, there exist bijections $f : X \to \omega$ and $g : Y \to \omega$. Then $f \times g : X \times Y \to \omega \times \omega$ is a bijection. Finally, we compose this with a bijection $h : \omega \times \omega \to \omega$ to give us a bijection $h \circ (f \times g) : X \times Y \to \omega$. □

Using our bijection $\omega \times \omega \cong \omega$, we can also prove that the countably-infinite union of countably-infinite sets is countably-infinite:

Theorem 4.8

Let X be a countably-infinite family of countable sets. Then $\bigcup_{x \in X} x$ is countably-infinite.

Proof.
This is the first theorem in which the Axiom of Choice (or rather, a weaker version) is required to prove the result, so we will be careful to point out exactly where its use comes in.

Let $f : \omega \to X$ be a bijection, and denote the image $f(n)$ by x_n. For each $n \in \omega$, we let F_n denote the set of all injections from x_n into ω. Because x_n is countable, each F_n is non-empty. We then have the family $\{F_n \mid n \in \omega\}$, which is a set by the Axiom of Replacement. Using the Axiom of Choice, we choose for every n an element $f_n \in F_n$. We then define a map $g : \bigcup_{x \in X} x \to \omega \times \omega$ defined by

$$g(y) := (n, f_n(y))$$

where n is the smallest natural number such that $y \in x_n$. This is an injection, as if $(n, f_n(y)) = (m, f_m(y'))$, then $n = m$ and thus $f_m(y') = f_n(y')$, and also $f_n(y) = f_n(y')$. But because f_n is an injection, we find that $y = y'$. Composing with a bijection $g : \omega \times \omega \to \omega$ gives us an injection of $\bigcup_{x \in X} x$ into ω, and thus $\bigcup_{x \in X} x$ is at-most countable.

If $\bigcup_{x \in X} x = U$ were finite, then each set x would be a subset of U, i.e. an element of $\mathcal{P}(U)$. But there are only finitely-many elements of U, whereas there are infinitely-many elements of X, giving us a contradiction. Thus, $\bigcup_{x \in X} x$ is in fact countably-infinite. $\square$

So far we have only seen countably-infinite sets, and shown that (finite) products of countably-infinite sets and the countably-infinite union of countably-infinite sets are countably-infinite, so we might be persuaded to believe that all infinite sets are countably-infinite. However, we can easily produce an example of an uncountable set by making use of the power set:

Theorem 4.9: Cantor's Theorem

There exist uncountable sets. Namely, $\mathcal{P}(\omega)$ is uncountable, and more generally, $|S| < |\mathcal{P}(S)|$ for all sets S.

Proof.
We shall use a methodology known as **Cantor's Diagonal Argument** which can essentially be thought of as showing that an enumeration of objects isn't surjective by defining an object that disagrees with the enumeration somewhere. To give some explanation to this and to provide some intuition for the abstract statement, we work informally through a simple example. Recall that $\mathcal{P}(S) \cong 2^S$ for every set S, where a function $f : S \to \{0, 1\}$ is mapped to the subset consisting of the pre-image of 1; in this sense, '$f(s) = 1$' indicates that the element s is in the subset (corresponding to f) and '$f(s) = 0$' indicates that the element s is not in the subset (corresponding to f).

Thus, we can represent every subset of ω by a sequence of 0's and 1's. Suppose that $\mathcal{P}(\omega) \cong 2^\omega$ were countably-infinite; then there would be an enumeration of all of the elements of 2^ω. Let's represent the beginning of such an enumeration in the following table:

	0	1	2	3	4	5	$\cdots$
s_0	0	1	1	0	1	0	$\cdots$
s_1	1	0	1	1	1	1	$\cdots$
s_2	0	1	1	1	0	1	$\cdots$
s_3	1	0	0	1	1	0	$\cdots$
s_4	1	1	0	0	0	0	$\cdots$
s_5	0	1	0	1	0	1	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$

Then we define a new element s such that

$$s(i) := \begin{cases} 0 & s_i(i) = 1 \\ 1 & s_i(i) = 0 \end{cases}.$$

That is, we look at the diagonal of this table, and exchange 0's for 1's and vice-a-versa. So, with the above example enumeration, we have $s(0) = 1, s(1) = 1, s(2) = 0, s(3) = 0, s(4) = 1, s(5) = 0, \ldots$. Then we ask if s is the image of this enumeration; if so, then $s = s_n$ for some n. But by definition $s(n) \neq s_n(n)$, meaning that s is not in the image, and thus the enumeration is not complete. In particular, no injection is ever a surjection, showing us that $|\omega| < |2^\omega|$.

To generalize to any set S, suppose that we have an injection $f : S \to \mathcal{P}(S)$. We define the subset

$$T := \{x \in S \mid x \notin f(x)\}.$$

Now, if T is in the image of f, then there exists $s \in S$ such that $f(s) = T$. Next, we derive a contradiction by considering the question: $s \in T$?. If $s \in T$, then by definition $s \notin f(s) = T$, giving us a contradiction. On the other hand, if $s \notin T$, then by the definition of T we have $s \in T = f(s)$, giving us another contradiction. Since either $s \in T$ or $s \notin T$, we must conclude that s not exist, and thus T does not lie in the image. This shows us that f is not a surjection.

To see that there exist any injections, take the simple injection $f : S \to \mathcal{P}(S)$ defined by $f(s) = \{s\}$. This shows us that $|S| \leq |\mathcal{P}(S)|$, where as the above forces this to be $|S| < |\mathcal{P}(S)|$.

$\square$

Section 5

Ordinals, Cardinals, and the Axiom of Choice

In this final section we look at implications of the Axiom of Choice, including that $|A| \leq |B|$ or $|B| \leq |A|$ for every pair of sets A, B, that every set is well-orderable, and a result known as Zorn's Lemma, which is the most complicated of the implications we will look at but one used periodically.

In order to get to this point, we look at a particular class of well-ordered sets called ordinals; these are the canonical examples of well-ordered sets. From here, we establish the fact that every well-ordered set corresponds to exactly one ordinal, backing up the claim that they are canonical. Moreover, these ordinals include a class known as cardinals which act as canonical representatives of size.

Subsection 5.1

Ordinals

An **ordinal** is a set x that is

(i) strictly well-ordered by the membership relation $\in$ and

(ii) satisfies the property that given $y \in x$, then $y \subset x$, called *transitivity*.

It should be noted that the latter transitivity is not the same as the transitivity of $\in$ guaranteed by the fact that x is strictly well-ordered by $\in$ for this latter transitivity says that if $y \in x$ and $z \in y$, then $z \in x$, while the transitivity of $\in$ as a strict well-order on x says given $w \in z$, $z \in y$, then $w \in y$: the difference is that our latter transitivity says something about x, while the transitivity of the order $\in$ says something about *the elements in* x. A set x that satisfies this latter transitivity is called a **transitive set**.

The following proposition and its corollaries show us some examples of ordinals, including many familiar examples:

Proposition 5.1

$\varnothing$ is an ordinal, for each ordinal α, $\alpha^+ = \alpha \cup \{\alpha\}$ is an ordinal, and ω is an ordinal.

Proof.
$\varnothing$ is trivially an ordinal as the membership relation on $\varnothing$ is trivial, and thus certainly is a strict well-order. That $\varnothing$ is a transitive set follows trivially as well, since there are no elements of $\varnothing$ to consider.

Now suppose that α is an ordinal. Then $\alpha^+ = \alpha \cup \{\alpha\}$ is strictly totally ordered by $\in$. To see irreflexitivity, suppose that $\beta \in \alpha^+$. If $\beta \in \alpha$, then by the fact that the relation $\in$ restricted to α is irreflexive, we find that $\beta \notin \beta$. On the other hand, if $\beta = \alpha$ and $\alpha \in \alpha$, then because α is an element of α it follows from the irreflexitivity of $\in$ restricted to α that $\alpha \notin \alpha$, giving us a contradiction. Thus, it must be the case that $\alpha \notin \alpha$, and irreflexitivity of $\in$ over α^+ follows. Because every element of α lies in α, it follows that α is the greatest element of α^+. To see that transitivity holds, suppose $x \in y \in z$. If $x, y, z \in \alpha$, then this holds by restriction of $\in$ to α. On the other hand, if any of x, y, z are α, then the only possibility is $z = \alpha$ since α is the maximum element and $\in$ is irreflexive. Then $x, y \in \alpha$ from which it follows that $x \in z$ automatically. This shows us that α^+ is strict totally ordered by $\in$.

To see that it is a well-ordered, let S be a non-empty subset of α^+. If $S \smallsetminus \{\alpha\}$ is non-empty, then it is a non-empty subset of α and thus has a least element β. Since $\beta \in \alpha$ automatically, it follows that $\beta \in \gamma$ for all $\gamma \in S \smallsetminus \{\beta\}$ and is thus a minimum. On the other hand, if $S \smallsetminus \{\alpha\} = \varnothing$, then $S = \{\alpha\}$ and it follows that α is the least element. Thus, α^+ is well-ordered.

It only remains to show that α^+ is a transitive set. Suppose that $x \in \alpha^+$. Then either $x = \alpha$ or $x \in \alpha$. In the former case, $x = \alpha \subset \alpha^+$, and in the latter case $x \subset \alpha$ by the fact that α is an ordinal, and thus $x \subset \alpha^+$ by transitivity of the subset relation. This shows that α^+ is an ordinal.

To see that ω is an ordinal, recall from **??** that ω is strictly well-ordered by $\in$, and from **??** we see that ω is a transitive set. □

Corollary 5.2

For every ordinal α, $\alpha \notin \alpha$.

Proof.
This was proved above by considering if $\alpha \in \alpha$, then because α is strictly well-ordered by $\in$ it follows that $\alpha \notin \alpha$, giving us a contradiction. □

The next corollary gives us an infinitude of examples of ordinals:

Corollary 5.3

Every natural number is an ordinal.

Proof.
Let

$$S := \{n \in \omega \mid n \text{ is an ordinal}\}.$$

We know that $\varnothing \in S$ by the above proposition, and if $n \in S$, then n^+ is an ordinal as well, so that S is inductive and $S = \omega$. □

We say that an ordinal α is a **successor ordinal** if there is an ordinal β such that $\beta^+ = \beta \cup \{\beta\} = \alpha$. If an ordinal α is not a successor ordinal, then it is said to be a **limit ordinal**.

We can define an ordering $<$ between ordinals by declaring $\alpha < \beta$ if and only if $\alpha \in \beta$. Many of these properties will be reminiscent of the order properties on the natural numbers, which should be expected since the natural numbers are ordinals themselves:

Lemma 5.4

If α is an ordinal, then $\alpha = 0$ or $\alpha > 0$.

Proof.
Suppose $\alpha \neq 0$. Then α is non-empty, and there exists a least element $\beta \in \alpha$. Because α is transitive, it follows that $\beta \subset \alpha$. If β is non-empty, then there is $\gamma \in \beta$. But then $\gamma \in \beta \subset \alpha$ so $\gamma \in \beta$. But $\gamma \in \beta$, contradicting the minimality of β. It thus follows that $\beta = \varnothing$ so that $0 < \alpha$. □

Lemma 5.5

If α, β are ordinals, then $\alpha \leq \beta$ if and only if $\alpha \subset \beta$.

Proof.
Suppose $\alpha \leq \beta$. Then either $\alpha \in \beta$ or $\alpha = \beta$. In the first case, by the fact that β is transitive we find that $\alpha \subset \beta$, and in the second case it is obvious that $\alpha \subset \beta$.

Conversely, suppose $\alpha \subsetneq \beta$. Thus, $\beta \setminus \alpha$ is non-empty, so has a least element x. We wish to show that $x = \alpha$, establishing that $\alpha \in \beta$. To do this, we show that $x \subset \alpha$ and $\alpha \subset x$.

$x \subset \alpha$: By the minimality of x, if $y \in x$, then $y \notin \beta \setminus \alpha$. Thus, we find that

$$\varnothing = x \cap (\beta \setminus \alpha) = (x \cap \beta) \setminus \alpha = x \setminus \alpha$$

making use of the transitivity of β to reduce $x \cap \beta$ to x. Because $x \smallsetminus \alpha = \varnothing$, we see that $x \subset \alpha$.

$\alpha \subset x$: Suppose that $y \in \alpha$. Because $\alpha \subset \beta$, we find that $y \in \beta$, and so by applying trichotomy of $\in$ for β, we see that exactly one of $x \in y, x = y, y \in x$ holds. If either of $x \in y$ or $x = y$ hold, then we find (applying the transitivity of α in the former case) that $x \in \alpha$, contradicting the fact that $x \in \beta \smallsetminus \alpha$. Thus, the only possibility is that $y \in x$ for each $y \in \alpha$, i.e. that $\alpha \subset x$.

Thus, we find that $\alpha = x$ and so $\alpha \in \beta$. □

Corollary 5.6

If α and β are ordinals and $\alpha < \beta$, then $\alpha^+ \leq \beta$.

Proof.
We know that $\alpha < \beta$ means $\alpha \in \beta$, so that by transitivity we have $\alpha \subset \beta$. But $\alpha \in \beta$ and $\alpha \subset \beta$ implies $\alpha^+ \subset \beta$ by **??**. Thus, $\alpha^+ \leq \beta$. □

In particular, we can show that this is a strict well-order on any set of ordinals. We begin with the following lemma:

Lemma 5.7

If α and β are ordinals, then $\alpha \cap \beta$ is an ordinal.

Proof.
Because $\alpha \cap \beta \subset \alpha$, it follows that $\alpha \cap \beta$ is well-ordered by $\in$. To see that it is transitive, suppose that $x \in \alpha \cap \beta$. Then $x \in \alpha$ and $x \in \beta$, implying $x \subset \alpha$ and $x \subset \beta$, which together implies $x \subset \alpha \cap \beta$. □

Proposition 5.8

Every set S of ordinals is strictly well-ordered by $<$.

Proof.
We must first show that $<$ is irreflexive, transitive, and trichotomous. Irreflexitivity has already been proven, and transitivity follows from the fact that if $\alpha < \beta$ and $\beta < \gamma$, then $\alpha \in \beta$ and $\beta \subset \gamma$. Thus, $\alpha \in \gamma$. To show trichotomy, assume for the sake of a contradiction that we have two ordinals $\alpha \neq \beta$. If $\alpha < \beta$, then $\alpha \cap \beta = \beta$, and if $\beta < \alpha$, then $\alpha \cap \beta = \alpha$. If neither of these hold, then $\alpha \cap \beta \subsetneq \beta$ and $\alpha \cap \beta \subsetneq \alpha$. By **??** we see that $\alpha \cap \beta \in \alpha$ and $\alpha \cap \beta \in \beta$. But this implies that $\alpha \cap \beta \in \alpha \cap \beta$, which by **??** this gives us a contradiction.

It remains to show that S is strictly well-ordered by $<$. Let T be a non-empty subset of S. Because T is non-empty, let $\alpha \in T$. If $\alpha \cap T = \varnothing$, then there is no $\beta \in T$ such that $\beta < \alpha$, so α is the minimum of T. Otherwise, $\alpha \cap T \neq \varnothing$, and because this is a subset of α, there is a least element β of $\alpha \cap T$. We claim β is the minimum of T. Othewise, there is $\gamma \in T$ such that $\gamma < \beta$. But by transitivity of α, $\gamma \in \alpha$, so $\gamma \in \alpha \cap T$. But this contradicts our choice of β as the smallest element in $\alpha \cap T$. Thus, β is the minimum of T. This means that T always has a minimum, and so S is well-ordered by $<$. □

This can give an alternate proof that ω is an ordinal, but it is also important in proving the existence of other ordinals aside from the natural numbers and ω.

In particular, every ordinal is a set of ordinals, namely every ordinal less than it:

LEMMA 5.9

Every element of an ordinal is an ordinal.

Proof.
Let α be an ordinal and suppose $\beta \in \alpha$. Then $\beta \subset \alpha$. β is a subset of the well-ordered set α, so it is well-ordered by the same relation $\in$. To show that it is a transitive subset, suppose that $\gamma \in \beta$. Because $\beta \subset \alpha$, we have $\gamma \in \alpha$, and by the transitivity of α we have $\gamma \subset \alpha$. If we have $\delta \in \gamma$, then $\delta \in \alpha$, so that by the transitivity of $\in$ in α we find that $\delta \in \gamma \in \beta$, showing that $\gamma \subset \beta$. □

LEMMA 5.10

The union of any set of ordinals is an ordinal.

Proof.
Let S be any set of ordinals, and let $U = \bigcup_{\alpha \in S} \alpha$. Since every element of an ordinal is an ordinal by **??** it follows that U is well-ordered by $\in$. It remains to show that U is transitive. To see this, suppose that $\beta \in U$. Then there is $\alpha \in S$ such that $\beta \in \alpha$. Since α is an ordinal, we see that $\beta \subset \alpha$, and thus $\beta \subset \alpha \subset U$. □

PROPOSITION 5.11

Let S be any set of ordinals. Then

$$\bigcup_{\alpha \in S} \alpha$$

is the least ordinal greater than or equal to each element of S.

Proof.
By **??**, it follows that $\bigcup_{\alpha \in S} \alpha$ is an ordinal.

$\beta \leq \alpha$ if and only if $\beta \subset \alpha$ per **??**. Thus, $\beta \leq \bigcup_{\alpha \in S} \alpha$ for every $\beta \in S$, showing that $\bigcup_{\alpha \in S} \alpha$ is an upper bound of S.

To see that it is the least upper bound, suppose that there is an ordinal γ such that $\beta \leq \gamma \leq \bigcup_{\alpha \in S} \alpha$ for each $\beta \in S$. Then it must be that $\gamma = \bigcup_{\alpha \in S} \alpha$ since if $\gamma \neq \bigcup_{\alpha \in S} \alpha$, then there is $\delta \in \bigcup_{\alpha \in S} \alpha$ not in γ. Since $\delta \in \bigcup_{\alpha \in S} \alpha$, there is $\beta \in S$ such that $\delta \in S$, showing that $\beta \not\subset \gamma$, contrary to assumption. Thus, $\bigcup_{\alpha \in S} \alpha$ is the least upper bound. □

This provides a useful characterization of limit ordinals:

COROLLARY 5.12

An ordinal is a limit ordinal if and only if it equals the union of all ordinals less than it.

Proof.
Suppose that $\alpha = \bigcup_{\beta < \alpha} \beta$ and suppose $\alpha = \gamma^+$, so that α is a successor ordinal. Then $\gamma < \alpha$ and by **??** if $\beta < \alpha$, then $\beta \leq \gamma$. Thus, $\gamma = \bigcup_{\beta < \alpha} \beta < \alpha$. This shows that it is sufficient for an ordinal to be the union of all ordinals less than it in order to be a limit ordinal.

Conversely, suppose that α is a limit ordinal. Since $\beta < \alpha$ for every β such that $\beta < \alpha$ (after all, this is what the statement means), it follows that α is an upper bound of the set of all ordinals less than α, and by **??** it follows that $\bigcup_{\beta < \alpha} \beta \leq \alpha$. Then if $\bigcup_{\beta < \alpha} \beta < \alpha$ it follows that $\left(\bigcup_{\beta < \alpha} \beta\right)^+ < \alpha$ since α is not a successor ordinal and by **??**. It follows that $\left(\bigcup_{\beta < \alpha} \beta\right)^+ \leq \bigcup_{\beta < \alpha} \beta$ since $\bigcup_{\beta < \alpha} \beta$ is an upper bound for all ordinals less than α. However, this contradicts the fact that $\gamma < \gamma^+$ for every ordinal γ. Thus, we conclude that $\alpha = \bigcup_{\beta < \alpha} \beta$. □

Subsection 5.2

Well-Ordered Sets and Ordinals

We the basic facts concerning ordinals under our belt, the most important of which that they are strict totally-ordered by the membership relation $\in$, we can prove several useful facts concerning general well-ordered sets, namely that ordinals act as canonical representatives of well-orderings.

First we recall the concept of downwards closure: a downwards closed subset T of a well-ordered set S is a subset of S such that if $t \in T$ and $t' < t$, then $t' \in T$. By **??** we know that these are exactly the initial segments of S, $S{\downarrow}_x = \{y \in S \mid y < x\}$, and S itself. Because of the fact every element of an ordinal is an ordinal (**??**) and that the ordinals are totally ordered by $\in$ it follows that the initial segment $\alpha{\downarrow}_\beta$ for $\beta \in \alpha$ is simply β. When referring to the initial segment of an ordinal, we shall tend to use the notation $\alpha{\downarrow}_\beta$ if only to stress that we are considering β as the set of all ordinals less than β and as β when we are treating it as an ordinal.

To show that ordinals are canonical representatives of well-ordered sets, we must show that that for every well-ordered S *there exists* a *unique* ordinal that is order isomorphic to S. Thus, we must show that not only does there exist such an ordinal that is isomorphic to S, but that that ordinal is unique. It will be enough to show that no two distinct ordinals α, β are order isomorphic, as we could take two order isomorphisms $f : \alpha \to S$ and $g : \beta \to S$ to produce an order isomorphism $g^{-1} \circ f : \alpha \to \beta$ and derive a contradiction. To prove this, we prove a more general statement regarding initial segments of a well-ordered set, using the knowledge that given any two ordinals then one is an initial segment of the other to provide motivation.

For brevity, we shall use $\cong$ to denote the existence of an order isomorphism.

Proposition 5.13

If S is well-ordered and $x \in S$, then $S \ncong S{\downarrow}_x$.

Proof.
Suppose $f : S \to S{\downarrow}_x$ is an order isomorphism, and let $T = \{y \in S \mid f(y) \neq y\}$. If $T \neq \varnothing$, then $f(x) = x$ and thus $x \in S{\downarrow}_x$, giving us a contradiction. Otherwise, let y be the least element of T, so that because y is the least element of $S \setminus S{\downarrow}_y$, we know that $f(y)$ is the least element of $f[S] \setminus f[S{\downarrow}_y]$ since f is an order isomorphism. But $f[S] = S{\downarrow}_x$ and $S{\downarrow}_y$ is fixed by f by the definition of y. Thus, $f(y)$ is the least element of $S{\downarrow}_x \setminus S{\downarrow}_y$. But if this is non-empty, then y is the least element of $S{\downarrow}_x \setminus S{\downarrow}_y$ and thus $f(y) = y$, contradicting the fact that $y \in T$. It it is empty, then there cannot exist y. In either case, we have a contradiction and f cannot exist. □

Corollary 5.14

Two distinct ordinals cannot be order isomorphic.

Proof.
If α and β are two ordinals such that $\alpha \neq \beta$, then we can assume without loss of generality that $\alpha < \beta$, so that $\alpha = \beta{\downarrow}_\alpha$. But by **??** we see that it is not possible for $\alpha \cong \beta$ to be true. □

Corollary 5.15

If S is a well-ordered set and $x, y \in S$, then $S{\downarrow}_x \cong S{\downarrow}_y$ if and only if $x = y$.

Proof.
If $x = y$ then it is obvious that $S{\downarrow}_x \cong S{\downarrow}_y$ (they are in fact equal).

For the converse, we can assume for the sake of the contradiction that $x \neq y$ and $S{\downarrow}_x \cong S{\downarrow}_y$. Without loss of generality we assume $x < y$. But then $S{\downarrow}_x \subset S{\downarrow}_y$ and $S{\downarrow}_x$ is downwards closed in $S{\downarrow}_y$. Thus, $(S{\downarrow}_y)_x = S{\downarrow}_x$, and we see from **??** that $(S{\downarrow}_y)_x = S{\downarrow}_x \ncong S{\downarrow}_y$, giving us a contradiction. Thus, $x = y$. □

To show the existence of an ordinal order isomorphic to a given well-ordered set, we prove a more general fact:

Theorem 5.16

Let S and T be well-ordered sets. Then $S \cong T$, $S{\downarrow_x} \cong T$ for some $x \in S$, or $S \cong T{\downarrow_y}$ for some $y \in T$.

Proof.
We shall define a subset of $S \times T$ by

$$f = \{(x, y) \in S \times T \mid S{\downarrow_x} \cong T{\downarrow_y}\}$$

which we must show is a function over some domain. To see that it is well-defined, suppose we have $(x, y_1), (x, y_2) \in f$. Then $T{\downarrow_{y_1}} \cong T{\downarrow_{y_2}}$, which by **??** we conclude that $y_1 = y_2$. Thus, f is well-defined and is a function over some subset of S. f is an injection as well, as $f(x_1) = f(x_2)$ implies $S{\downarrow_{x_1}} \cong S{\downarrow_{x_2}}$ which by **??** we conclude that $x_1 = x_2$. Finally, if $x_1 < x_2$, then $T{\downarrow_{f(x_1)}}$ is isomorphic to an initial segment of $T{\downarrow_{f(x_2)}}$ because they are isomorphic to $S{\downarrow_{x_1}}$ and $S{\downarrow_{x_2}}$. Thus, $f(x_1) < f(x_2)$ because if $f(x_1) > f(x_2)$ then $T{\downarrow_{f(x_1)}}$ would be isomorphic to one of its initial segments. Thus, we conclude that f is an order isomorphism from its domain to its image.

The domain and image of f must be downwards closed, as if $S{\downarrow_x} \cong T{\downarrow_y}$, then every initial segment of $S{\downarrow_x}$ is isomorphic to an initial segment of $T{\downarrow_y}$ by restricting the order isomorphism, and every initial segment of $T{\downarrow_y}$ is isomorphic to an initial segment of $S{\downarrow_x}$ by doing similarly. Thus, the domain is either $S{\downarrow_x}$ for some $x \in S$ or S, and the image is either $T{\downarrow_y}$ for some $y \in T$ or T. If the domain is unequal to S and the image is unequal to T, then we let x be the least element of S not in the domain, and similarly y the least element of T not in the image. But then we can extend f to x by mapping x to y, giving us a contradiction that $(x, y) \notin f$. Thus, either S is the domain or T is the image (or both). ∎

Corollary 5.17

Every well-ordered set is order isomorphic to a unique ordinal.

Proof.
Let S be a well-ordered set, and define

$$T = \{x \in S \mid \text{there exists an ordinal } \alpha \text{ such that } \alpha \cong S{\downarrow_x}\}.$$

For each $x \in T$ the ordinal α with $\alpha \cong S{\downarrow_x}$ is unique by **??**. Thus, by using the Axiom of Replacement we have that the collection

$$\{\alpha \mid \alpha \text{ is an ordinal and there is } x \in S \text{ with } S{\downarrow_x} \cong \alpha\}$$

is a set, which we call β. β is an ordinal, as it is well-ordered by $\in$ since it is a set of ordinals (**??**), and that it is a transitive set follows because if $\alpha \in \beta$ and $\gamma \in \alpha$, then there is an order isomorphism $f : \alpha \to S{\downarrow_x}$ for some $x \in S$ and then

$$\gamma = \alpha{\downarrow_\gamma} \cong (S{\downarrow_x})_{f(y)} = S{\downarrow_{f(y)}}$$

so that $\gamma \in \beta$.

Now, by **??** either $\beta \cong S{\downarrow_x}$ for some $x \in S$, $\beta{\downarrow_\alpha} \cong S$ for some $\alpha \in \beta$, or $\beta \cong S$. In the first case, $\beta \in \beta$, so that by **??** we reach a contradiction. In the second case, $\beta{\downarrow_\alpha} = \alpha \cong S$, and in the third case $\beta \cong S$, so that in either case S is isomorphic to an ordinal. That this ordinal is unique follows from **??**. ∎

This shows us that, indeed, there exists a unique ordinal isomorphic to a given well-ordered set, giving credibility to the claim that the ordinals act as canonical representatives of well-ordered sets. Given a well-ordered set S, we denote by $\mathrm{type}(S)$ the unique ordinal order isomorphic to S. Once we prove that every set can be well-ordered, this fact will be a useful tool.

Subsection 5.3

Implications of the Axiom of Choice

The first implication of the Axiom of Choice that we will deal with is the **Well-Ordering Theorem**, which states that every set can be well-ordered. Essentially what we'd like to do is use a choice function to pick out of non-empty subsets a least element, defining a bijection with an ordinal recursively. However, an issue arises: given an ordinal α, how do we know we won't run out of elements in α during this process? To ensure that we can pick an ordinal so that this doesn't happen, we prove that for every set there is an ordinal 'bigger' than that set:

Proposition 5.18: Hartogs Number

Given any set S, there exists an ordinal α such that there is no injection from α to S.

Proof.
Let

$$\alpha = \{\beta \mid \beta \text{ is an ordinal and there is an injective map from } \beta \text{ to } S\}$$

α is an ordinal, as it is a set of ordinals and thus by **??** is well-ordered by $\in$, and it is transitive as given $\beta \in \alpha$ and $\gamma \in \beta$ there is an injection $f : \beta \to S$ we can restrict this to an injection $f|_\gamma : \gamma \to S$, showing that $\gamma \in \alpha$.

Now, $\alpha \notin \alpha$ by **??**, and thus there is no injection from α into S. □

We this result, we can proceed with our proof:

Theorem 5.19: Well-Ordering Theorem

Every set can be well-ordered.

Coupled with **??**, we can use the Well-Ordering Theorem to show that any two sets are comparable in the sense that for any two sets S and T there is either an injection from S into T or T into S.

Theorem 5.20: Trichotomy

Given any two sets S and T, either $|S| < |T|$, $|S| = |T|$, or $|T| < |S|$.

Proof.
We begin by well-ordering S and T, possible by **??**. Then by **??** we know that exactly one of $S{\downarrow_x} \cong T$ for some $x \in S$, $S \cong T{\downarrow_y}$ for some $y \in T$, or $S \cong T$ holds. These corresponds to $|T| < |S|$, $|S| < |T|$, or $|S| = |T|$. □

Subsection 5.4

Cardinals and Cardinality

We now end by making the notation $|X|$ more than a formalism. An ordinal that is not in bijection with any smaller ordinal is called a **cardinal**. While ordinals are canonical representatives of well-orders, the cardinals are the canonical representatives of size.

Some simple examples of cardinals include the natural numbers, which follows from **??**. Similarly, ω is a cardinal because of the fact that it is infinite, and every ordinal less than it is finite (since they are natural numbers). **??** guarantees that given a cardinal, we can always create a cardinal greater than it. When treating ω as a cardinal, it is common to denote it by $\aleph_0$, where $\aleph$ is the Hebrew letter 'aleph'.

Lemma 5.21

Every set is in bijection with a unique cardinal.

Proof.
Let S be any set. Using the Well-Ordering Theorem, we well-order S. Then by **??** we know that there exists an ordinal α such that S and α are order isomorphic. In particular, the order isomorphism is a bijection. We then define the set

$$T = \{\beta \in \alpha^+ \mid \beta \text{ is an ordinal such that } |\beta| = |S|\}$$

and let κ be the least element of T. Then κ is a cardinal because if γ was an ordinal such that $|\kappa| = |\gamma|$, then $|\gamma| = |S|$ and γ is an element of T. But since κ is the least element of T, it follows that $\kappa \leq \gamma$, and thus κ is a cardinal.

Uniqueness follows from the fact that if λ were another cardinal in bijection with S, then λ and κ would be in bijection. But because they are both cardinals, we would have $\kappa \leq \lambda$ and $\lambda \leq \kappa$, from which we conclude that $\kappa = \lambda$. □

This shows us that the cardinals act as canonical representatives of size: the **cardinality** of a set S, denoted $|S|$, is the unique cardinal in bijection with S.

We now show that the formal inequalities $|S| \leq |T|$ correspond exactly with the ordering on the ordinals when restricted to the cardinals.

Lemma 5.22

Let κ and λ be cardinals. Then $\kappa \leq \lambda$ as ordinals if and only if there is an injection from κ to λ, with $\kappa = \lambda$ if and only if there is a bijection between them.

Proof.
If $\kappa \leq \lambda$, then $\kappa \subset \lambda$ and thus there is the inclusion of κ into λ as an injection. Conversely, if there is an injection of κ into λ, then it cannot be the case that $\kappa > \lambda$, as then there would be an injection of λ into κ, from which by the Cantor-Schroeder-Bernstein Theorem shows that there is a bijection between κ and λ. But this would contradict the fact that κ is a cardinal, as there would be an ordinal in bijection with it that was strictly less than it. But, it must be that $\kappa \leq \lambda$.

Equality if and only if there is a bijection follows directly from the definition of being a cardinal: if $\kappa = \lambda$, then the identity function is such a bijection, and if there is a bijection between κ and λ, since both are cardinals we have $\kappa \leq \lambda$ and $\lambda \leq \kappa$, leading us to $\kappa = \lambda$. □

Corollary 5.23

For all sets S and T, $|S| \leq |T|$ if and only if there is an injection from S into T, and $|S| = |T|$ if and only if there is a bijection between S and T.

Since the ordering on cardinals is exactly the ordering on the ordinals when restricted to the cardinals, it follows that every set of cardinals is well-ordered by $\in$.

Lemma 5.24

There is no set of all ordinals, cardinals, or infinite cardinals.

Proof.
Suppose there existed a set $\mathcal{O}$ of all ordinals. Every set of ordinals is well-ordered, and more-over $\mathcal{O}$ is transitive since every element of an ordinal is another ordinal and thus an element of $\mathcal{O}$. This shows that $\mathcal{O}$ is moreover an ordinal itself. But then $\mathcal{O} \in \mathcal{O}$, contradicting **??**. Thus, there cannot exist such a set.

Now suppose there existed a set $\mathcal{C}$ of all cardinals. For every ordinal α, there is a larger cardinal, namely $|\mathcal{P}(\alpha)|$, and thus $\alpha \in |\mathcal{P}(\alpha)|$. Since $\mathcal{C}$ is a set, $\bigcup \mathcal{C}$ is also a set. But $\mathcal{O} \subset \bigcup \mathcal{C}$, and $\mathcal{O}$ is not a set, contradicting the existence of $\mathcal{C}$.

Finally, there can be no set of all infinite cardinals for the set of finite cardinals is precisely ω. If the set $\mathcal{I}$ of infinite cardinals existed, then $\mathcal{I} \cup \omega = \mathcal{C}$ would exist, leading to a contradiction. ☐

Theorem 5.25

For every ordinal α, there is an infinite cardinal κ such that the set

$$I_\kappa := \{\lambda \mid \lambda \text{ is an infinite cardinal and } \lambda < \kappa\}$$

of cardinals is order isomorphic to α.

Proof.
We begin by noting that I_κ is a set because every infinite cardinal λ less than κ is also an ordinal which is necessarily less than κ as an ordinal, and thus $\lambda \in \kappa$. For this reason, the set can be rewritten as $\{\lambda \in \kappa \mid \lambda \text{ is an infinite cardinal and } \lambda < \kappa\}$.

Now, suppose for the sake of a contradiction that there exists an ordinal α such that there is no κ such that α is order isomorphic to I_κ. If α is a successor ordinal, then there is an ordinal β such that $\beta^+ = \alpha$. By hypothesis, there exists an infinite cardinal λ such that β is order isomorphic to I_λ, say with order isomorphism $f : I_\lambda \to \beta$. Letting λ be the least cardinal greater than λ, we can form an order isomorphism between $I_\kappa = I_\lambda \cup \{\lambda\}$ and $\alpha = \beta^+ = \beta \cup \{\beta\}$ by mapping I_λ to β under f and sending λ to β. This leads to a contradiction, so α must be a limit ordinal.

Since α is a limit ordinal, it follows from **??** that $\alpha = \bigcup_{\beta<\alpha} \beta$. By hypothesis, for each $\beta < \alpha$ there exists an infinite cardinal λ_β such that I_{λ_β} is order isomorphic to β. Then $S = \bigcup_{\beta<\alpha} I_{\lambda_\beta}$ is a downwards-closed set of infinite cardinals which is order isomorphic to α since if $\operatorname{type}(S) > \alpha$, then α is order isomorphic to an initial segment of S. An initial segment of S is of the form $S{\downarrow}_\lambda = I_\lambda$ for some $\lambda \in S$. Since $S = \bigcup_{\beta<\alpha} I_{\lambda_\beta}$ and each I_{λ_β} is a downwards-closed set of infinite cardinals, it follows that $I_\lambda \subset I_{\lambda_\beta}$ for some $\beta < \alpha$. Since $\operatorname{type}(I_{\lambda_\beta}) < \alpha$, it follows that $\operatorname{type}(I_\lambda) < \alpha$, and thus $\alpha < \alpha$, giving a contradiction. On the other hand, if $\operatorname{type}(S) < \alpha$ then there is $\lambda_{\operatorname{type}(S)}$ such that $\operatorname{type}(S) = \operatorname{type}(I_{\lambda_{\operatorname{type}(S)}}) = \operatorname{type}(S{\downarrow}_{\lambda_{\operatorname{type}(S)}})$, leading to a contradiction of **??**. Thus, it must be that $\operatorname{type}(S) = \alpha$.

Now, by **??** it follows that $S = \bigcup_{\beta<\alpha} I_{\lambda_\beta}$ cannot contain all cardinals. Since S is a downwards closed set of infinite cardinals, there are infinite cardinals greater than any element of S; let κ be the least such infinite cardinal, so that $S = I_\kappa$. Then $\operatorname{type}(I_\kappa) = \alpha$, contrary to hypothesis. ☐

Given α an ordinal, we define $\aleph_\alpha$ to be the unique infinite cardinal such that α is order isomorphic to $\{\kappa \mid \kappa \text{ is an infinite cardinal and } \kappa < \aleph_\alpha\}$.

We end by looking at some of the ways that cardinality is affected by taking products or disjoint unions. Namely, it is possible to show, using the Axiom of Choice, that if κ and λ are cardinals and $\kappa \le \lambda$ where λ is infinite, then $|\kappa \coprod \lambda| = \lambda = |\kappa \times \lambda|$:

Theorem 5.26

If λ and κ are cardinals with $\kappa \le \lambda$, $\kappa \ne \varnothing$, and λ infinite, then $|\kappa \coprod \lambda| = \lambda = |\kappa \times \lambda|$.

Proof.
It suffices to show that $|\lambda \times \lambda| = \lambda$, for if this does hold then

$$\lambda \le \left|\kappa \coprod \lambda\right| \le \left|\lambda \coprod \lambda\right| \le |\lambda \times \lambda| = \lambda$$

where the necessary injections are $f_1 : \lambda \to \kappa \coprod \lambda$ given by $f_1 : \alpha \mapsto (\alpha, 1)$, $f_2 : \kappa \coprod \lambda \to \lambda \coprod \lambda$ given by $f_2 = \iota \coprod \mathrm{id}_\lambda$ where ι is the inclusion of κ into λ, and $|\lambda \times \lambda| = \lambda$ by hypothesis. This shows that $|\kappa \coprod \lambda| = \lambda$. Similarly

$$\lambda \le |\kappa \times \lambda| \le |\lambda \times \lambda| = \lambda$$

where the necessary injections are $g_1 : \lambda \to \kappa \times \lambda$ defined by $g_1 : \alpha \mapsto (\beta, \alpha)$ where β is a fixed element of κ, $g_2 : \kappa \times \lambda \to \lambda \times \lambda$ given by $g_2 = \iota \times \mathrm{id}_\lambda$ where ι is the inclusion of κ into λ, and $|\lambda \times \lambda| = \lambda$ by hypothesis. This shows that $|\kappa \times \lambda| = \lambda$.

Assume for the sake of a contradiction that there exist infinite cardinals μ such that $|\mu \times \mu| \ne \mu$. Let λ be the least such cardinal. We shall well-order $\lambda \times \lambda$, and it will be order isomorphic to a unique ordinal δ. Proving that $\delta \le \lambda$, it follows $\lambda \le |\lambda \times \lambda| = |\delta| \le |\lambda| = \lambda$ and hence $|\lambda \times \lambda| = \lambda$, giving us a contradiction and showing that $|\mu \times \mu| = \mu$ for all infinite cardinals μ.

We strictly well-order $\lambda \times \lambda$ by setting $(\alpha_1, \beta_1) \prec (\alpha_2, \beta_2)$ if and only if

1. $\max\{\alpha_1, \beta_1\} < \max\{\alpha_2, \beta_2\}$, or
2. $\max\{\alpha_1, \beta_1\} = \max\{\alpha_2, \beta_2\}$ and $\alpha_1 < \alpha_2$, or
3. $\max\{\alpha_1, \beta_1\} = \max\{\alpha_2, \beta_2\}$, $\alpha_1 = \alpha_2$, and $\beta_1 < \beta_2$

where $<$ is the strict well-ordering on λ (i.e. the membership relation $\in$). Naturally, it must be checked that this is a strict well-ordering, so first it is necessary to check that it is a strict total ordering at all, and then that every non-empty subset of $\lambda \times \lambda$ has a minimum.

To see that it is a strict total ordering, note that none of $\alpha < \alpha$, $\beta < \beta$, or $\max\{\alpha, \beta\} < \max\{\alpha, \beta\}$ since $<$ is a irreflexive, and thus $(\alpha, \beta) \nprec (\alpha, \beta)$ is irreflexive. Next, if $(\alpha_1, \beta_1) \prec (\alpha_2, \beta_2)$ and (α_2, β_2), then either

(a) $\max\{\alpha_1, \beta_1\} < \max\{\alpha_2, \beta_2\}$,

(b) $\max\{\alpha_1, \beta_1\} = \max\{\alpha_2, \beta_2\}$ and $\alpha_1 < \alpha_2$, or

(c) $\max\{\alpha_1, \beta_1\} = \max\{\alpha_2, \beta_2\}$, $\alpha_1 = \alpha_2$, and $\beta_1 < \beta_2$,

and similarly one of

(A) $\max\{\alpha_2, \beta_2\} < \max\{\alpha_3, \beta_3\}$,

(B) $\max\{\alpha_2, \beta_2\} = \max\{\alpha_3, \beta_3\}$ and $\alpha_2 < \alpha_3$, or

(C) $\max\{\alpha_2, \beta_2\} = \max\{\alpha_3, \beta_3\}$, $\alpha_2 = \alpha_3$, and $\beta_2 < \beta_3$

hold. We go through by a case-by-case analysis:

- If (a) and (A) hold, then by the transitivity of $<$ we find that $\max\{\alpha_1, \beta_1\} < \max\{\alpha_3, \beta_3\}$.
- If (b) and (B) hold, then by transitivity of $=$ and $<$ we find that $\max\{\alpha_1, \beta_1\} = \max\{\alpha_3, \beta_3\}$ and $\alpha_1 < \alpha_3$.
- If (c) and (C) hold, then by transitivity of $=$ and $<$ we find that $\max\{\alpha_1, \beta_1\} = \max\{\alpha_3, \beta_3\}$ and $\alpha_1 = \alpha_3$ and $\beta_1 < \beta_3$.

- If (a) and either (B) or (C) hold, then $\max\{\alpha_1,\beta_1\} < \max\{\alpha_2,\beta_2\} = \max\{\alpha_3,\beta_3\}$.
- If either (b) or (c) and (A) hold, then $\max\{\alpha_1,\beta_1\} = \max\{\alpha_2,\beta_2\} < \max\{\alpha_3,\beta_3\}$.
- If (b) and (C) hold, then by the transitivity of $=$ we find that $\max\{\alpha_1,\beta_1\} = \max\{\alpha_3,\beta_3\}$ and $\alpha_1 < \alpha_2 = \alpha_3$.
- If (c) and (B) hold, then $\max\{\alpha_1,\beta_1\} = \max\{\alpha_3,\beta_3\}$ and $\alpha_1 = \alpha_2 < \alpha_3$.

In every case, we find that $(\alpha_1,\beta_1) \prec (\alpha_3,\beta_3)$. Finally, to see that $\prec$ is total, note that given (α_1,β_1) and (α_2,β_2), either $\max\{\alpha_1,\beta_1\} < \max\{\alpha_2,\beta_2\}$ (so $(\alpha_1,\beta_1) \prec (\alpha_2,\beta_2)$), $\max\{\alpha_2,\beta_2\} < \max\{\alpha_1,\beta_1\}$ (so $(\alpha_2,\beta_2) \prec (\alpha_1,\beta_1)$), or $\max\{\alpha_1,\beta_1\} = \max\{\alpha_2,\beta_2\}$, in which case either $\alpha_1 < \alpha_2$ (and $(\alpha_1,\beta_1) \prec (\alpha_2,\beta_2)$), $\alpha_2 < \alpha_1$ (and $(\alpha_2,\beta_2) \prec (\alpha_1,\beta_1)$), or $\alpha_1 = \beta_1$, in which case either $\beta_1 < \beta_2$ (and $(\alpha_1,\beta_1) \prec (\alpha_2,\beta_2)$), $\beta_2 < \beta_1$ (and $(\alpha_2,\beta_2) \prec (\alpha_1,\beta_1)$), or $\beta_1 = \beta_2$ (and $(\alpha_1,\beta_1) = (\alpha_2,\beta_2)$). That this covers all the possible cases and that each case is distinct follows from the fact that $<$ is total.

To see that every non-empty subset S of $\lambda \times \lambda$ has a minimum, note that the set $T := \{\max\{\alpha,\beta\} \in \lambda \mid (\alpha,\beta) \in \lambda \times \lambda\}$ is a non-empty subset of λ and thus has a minimum γ under the strict well-order $\in$. The set $U := \{(\alpha,\beta) \in \lambda \times \lambda \mid \max\{\alpha,\beta\} = \gamma\}$ is a non-empty subset of $\lambda \times \lambda$, and thus $\pi_1[U]$ is a non-empty subset of λ. Then $\pi_1[U]$ has a minimum α^* under the strict well-order $\in$. Finally, the set $V := \{(\alpha^*,\beta) \in \lambda \times \lambda \mid \max\{\alpha^*,\beta\} = \gamma\}$ is a non-empty subset of $\lambda \times \lambda$, and thus $\pi_2[V]$ is a non-empty subset of λ. Then $\pi_2[V]$ has a minimum β^* under the strict well-order $\in$. The element (α^*,β^*) of S is, by construction, a minimum of S, for if $(\alpha,\beta) \in \lambda \times \lambda$ is distinct from (α^*,β^*), then either $\max\{\alpha^*,\beta^*\} < \max\{\alpha,\beta\}$ (and $(\alpha^*,\beta^*) \prec (\alpha,\beta)$) or $\max\{\alpha^*,\beta^*\} = \max\{\alpha,\beta\} = \gamma$, so that $(\alpha,\beta) \in U$; it is not possible that $\max\{\alpha,\beta\} < \max\{\alpha^*,\beta^*\}$ for this would contradict our choice of (α^*,β^*). Then either $\alpha^* < \alpha$ (and $(\alpha^*,\beta^*) \prec (\alpha,\beta)$) or $\alpha^* = \alpha$, so that $(\alpha,\beta) \in V$; it is not possible that $\alpha < \alpha^*$ for this would contradict our choice of α^*. Finally, if $\max\{\alpha^*,\beta^*\} = \max\{\alpha,\beta\}$ and $\alpha^* = \alpha$, it must be the case that $\beta^* < \beta$ for otherwise either $\beta < \beta^*$, contradicting our choice of β, or $\beta = \beta^*$, in which case $(\alpha,\beta) = (\alpha^*,\beta^*)$, contradicting the hypothesis that they be distinct.

Now that the relation $\prec$ has been shown to be a strict well-order of $\lambda \times \lambda$, we see that there exists a unique ordinal δ order isomorphic to $\lambda \times \lambda$. We must show that $\delta \leq \lambda$. We shall show this using a proof by contradiction: suppose that $\lambda < \delta$ for the sake of a contradiction. Thus, λ is an initial segment of $\delta = \lambda \times \lambda$. For this reason, there is an initial segment of $\lambda \times \lambda$ with cardinality λ. To achieve our contradiction, we shall show that every initial segment has cardinality strictly less than λ.

Every initial segment generated consisting of all elements less than (α,β) is contained in the initial segment consisting of all elements less than $(\max\{\alpha,\beta\},\max\{\alpha,\beta\})$ since $(\alpha,\beta) \preceq (\max\{\alpha,\beta\},\max\{\alpha,\beta\})$. Thus, it suffices to consider an initial segment consisting of all elements less than (α,α), which is precisely the set $(\alpha^+ \times \alpha^+) \setminus \{(\alpha,\alpha)\}$ since $(\beta,\gamma) \prec (\alpha,\alpha)$ if and only if $\max\{\beta,\gamma\} < \max\{\alpha,\alpha\} = \alpha$, $\max\{\beta,\gamma\} = \alpha$ and $\beta < \alpha$, or $\max\{\beta,\gamma\} = \alpha$ and $\beta = \alpha$ and $\gamma < \alpha$. The elements satisfying the first condition are precisely those such that $\beta,\gamma < \alpha$, the second those of the form (β,α) with $\beta < \alpha$, and the last those of the form (α,γ) with $\gamma < \alpha$; the union of these three sets is precisely $(\alpha^+ \times \alpha^+) \setminus \{(\alpha,\alpha)\}$. This set has size at most $|\alpha^+ \times \alpha^+|$.

If α is finite, then α^+ is finite and so is $\alpha^+ \times \alpha^+$. Thus, $|\alpha^+ \times \alpha^+| < \lambda$ since λ is infinite. On the other hand, if α is infinite, then $|\alpha^+| = |\alpha|$ because $\omega^+ \cong \omega$ (the sets ω and $\{\omega\}$ are countable and thus $\omega \cup \{\omega\}$ is also countable) and $\omega^+ \cong \omega \cup \{\alpha\}$ so that we can take the disjoint union of the bijection between $\omega \cup \{\alpha\}$ and ω and the identity on $\alpha \setminus \omega$. Thus $|\alpha^+ \times \alpha^+| = |\alpha \times \alpha|$ by taking the product of the bijection $f : \alpha^+ \to \alpha$. Since $\alpha < \lambda$ and λ is the least infinite cardinal such that $|\lambda \times \lambda| \neq \lambda$, we see that $|\alpha^+ \times \alpha^+| = |\alpha \times \alpha| = ||\alpha| \times |\alpha|| = |\alpha| < \lambda$. Thus, every initial segment of $\lambda \times \lambda$ has cardinality strictly less than λ, giving us a contradiction.

We thus find that $|\lambda \times \lambda| = \lambda$, contrary to the assumption that there exist infinite cardinals μ such that $|\mu \times \mu| \neq \mu$. □

A slightly easier (less thorough) version:

Theorem 5.27

For every $\alpha \in \mathrm{Ord}$, $\aleph_\alpha \aleph_\alpha = \aleph_\alpha$.

Proof.
We prove by induction on α. Without loss of generality, we shall assume $\alpha > 0$ (since we know that $\aleph_0 \aleph_0 = \aleph_0$).

We want to show that $\omega_\alpha \times \omega_\alpha$ bijects with ω_α. We shall well-order $\omega_\alpha \times \omega_\alpha$ by 'going up in squares'. That is, we define $(x, y) < (x', y')$ if either $\max\{x, y\} < \max\{x', y'\}$ or $\max\{x, y\} = \max\{x', y'\} = \beta$ and $y' = \beta, y < \beta$, or $x = x' = \beta, y < y'$, or $y = y' = \beta, x < x'$ (i.e. if the maxima are equal, then we move to the dictionary ordering with the significance given to the second coordinate, as with the well-ordering on the product of ordinals).

For any proper initial segment $I_{(x,y)}$, we have $I_{(x,y)} \subset \beta \times \beta$ for some $\beta < \omega_\alpha$ (as ω_α is a limit). Without loss of generality, β is infinite (we can always choose β to be infinite by the assumption that $\alpha > 0$). By our induction hypothesis, $\mathrm{card}(\beta \times \beta) = \beta < \mathrm{card}(\omega_\alpha)$. Hence, $I_{(x,y)}$ has order-type strictly less than ω_α.

Thus, the order-type of our well-ordering on $\omega_\alpha \times \omega_\alpha$ is less than or equal to ω_α, and in particular $\omega_\alpha \times \omega_\alpha$ injects into ω_α. Thus, we find that $\omega_\alpha \times \omega_\alpha \leftrightarrow \omega_\alpha$ (by Cantor-Schroeder-Bernstein). $\square$

Corollary 5.28

Let $\alpha \leq \beta$. Then $\aleph_\alpha + \aleph_\beta = \aleph_\alpha \aleph_\beta = \aleph_\beta$.

Proof.
We have

$$\aleph_\beta \leq \aleph_\alpha + \aleph_\beta \leq \aleph_\beta + \aleph_\beta = 2\aleph_\beta \leq \aleph_\alpha \aleph_\beta \leq \aleph_\beta \aleph_\beta = \aleph_\beta,$$

and hence we have equality throughout. $\square$