ORDINALS – A CRASH COURSE

Dominic van der Zypen

This self-explaining[1] note aims to give a concise yet thorough introduction to well-ordered sets and ordinals. All the basic concepts and propositions are given, and in the last chapter we prove that ordinals are prototypes for well-ordered sets. Moreover, an introduction to the Axiom of Choice is given and we define cardinals and cardinality. However, the note does not deal with more advanced topics such as ordinal and cardinal arithmetic or singularity vs regularity.

1 Well-ordered sets

A *relation* on a set X is a subset $R \subseteq X \times X$. It is very common to use infix notation, that is instead of $(x, y) \in R$ we write xRy. A *partially ordered set* or *poset* is a pair $(X, \leq)$ consisting of a set X and a relation $\leq$ on X such that for all $x, y, z \in X$

1. $x \leq x$;
2. $x \leq y$ and $y \leq x$ imply $x = y$;
3. $x \leq y$ and $y \leq z$ imply $x \leq z$.

The relation $\leq$ is referred to as the *ordering relation* of the corresponding poset. The statement "$x < y$ is by definition equivalent to "$x \leq y$ and $x \neq y$. Sometimes we are sloppy and write X instead of $(X, \leq)$ when the ordering relation $\leq$ is obvious.

It is well possible for elements x, y of a poset that $x \not\leq y$ and $y \not\leq x$. A *toset (or totally ordered set)* is a poset X such that for all $x, y \in X$ we have $x \leq y$ or $y \leq x$.

Let $(X, \leq_X)$ and $(Y, \leq_Y)$ be posets. A map $f : X \to Y$ is said to be *order-preserving* if $a \leq_X b$ implies $f(a) \leq_Y f(b)$. Further, f is an *order-isomorphism* if it is order-preserving, bijective and also its inverse f^{-1} is

[1] if you have had exposure to naive set theory and its notation such as it is offered in any introductory course to Linear Algebra or Analysis

order-preserving. In that case, X and Y are said to be *order-isomorphic*. If X, Y are totally ordered then the following holds: if $f : X \to Y$ is bijective and order-preserving, then it is an order-isomorphism (exercise: provide an example showing that this is not true in general for posets).

Moreover, we need the concept of the induced order. If $(X, \leq)$ is a poset and $S \subseteq X$, then $\leq \cap (S \times S)$ is an ordering relation on S which we refer to as the *induced ordering on S*. When talking about a subset of a poset we tacitly assume it to be endowed with the induced ordering.

A toset X is said to be *well-ordered* or a *woset* if every nonempty subset $Y \subseteq X$ has a smallest element (that is, there is $y_0 \in Y$ such that for all $y \in Y$ we have $y_0 \leq y$). Note that smallest elements are unique.

Subsets of a woset X are always wosets with the induced order. For any $a \in X$ we define

$$S_X(a) = \{x \in X : x < a\}.$$

Subsets of this form are called *(initial) segments*.

The following proposition (and its corollary) is crucial when dealing with wosets and later with ordinals.

PROPOSITION 1.1. *Let X be a well-ordered set. Suppose that $f : X \to Y$ is injective and order-preserving. Then for all $x \in X$ we have $x \leq f(x)$.*

Proof. Suppose this is not true. Then the set $M = \{x \in X : f(x) < x\}$ is non-empty and therefore has a smallest element x_0. But then we have $f(f(x_0)) \leq f(x_0)$ since f is order-preserving and $f(f(x_0)) \neq f(x_0)$ because f is injective. Therefore $f(x_0)$ is a member of M strictly smaller than x_0 contradicting x_0's minimality. □

COROLLARY 1.2. *Let X be a woset and $a \in X$. Then there is no order-isomorphism $f : X \to S_X(a)$.*

Proof. Suppose otherwise. Let $\iota : S_X(a) \to X$ be the inclusion map; so $\iota \circ f : X \to X$ is injective and order-preserving, but we have $f(a) \in S_X(a)$ and therefore $f(a) < a$, contradicting the proposition. □

The following proposition tells us when the union of well-ordered sets is again well-ordered:

PROPOSITION 1.3. *Let $\mathcal{C}$ be a collection of well-ordered subsets of a poset $(P, \leq)$. Suppose that for any two members X, Y of $\mathcal{C}$ either X is a segment of Y or the other way round. Then $\bigcup \mathcal{C}$ is well-ordered subset of P (with respect to the induced ordering).*

Proof. It is clear that $D := \bigcup \mathcal{C}$ is totally ordered in P. Let $\emptyset \neq A \subseteq D$. Show that A has a smallest element in D. There is $X \in \mathcal{C}$ such that $X \cap A \neq \emptyset$. So there is a smallest element $a_0 \in A \cap X$ in X since X is well-ordered. We claim that a_0 is the smallest element of $A \subseteq D$. Suppose not. Then let $a \in A$ be such that $a < a_0$. There is $Y \in \mathcal{C}$ such that $a \in Y$. But since $a \neq X$ we have that neither X is a segment of Y nor the other way round, contradiction. $\square$

2 Ordinals

DEFINITION 2.1. *An* ordinal *is a well-ordered set* $(X, \leq)$ *such that for all $a \in X$ we have $S_X(a) = a$.*

So, an ordinal is a well-ordered set such that each element **equals** the set of all its predecessors! Note that the condition "$S_X(a) = a$" in the definition directly implies the following observation:

For members α, β of an ordinal we have

$$\alpha < \beta \text{ if and only if } \alpha \in \beta.$$

The importance of this observation can not be overemphasised. It says that the ordering relation on any ordinal is given through the membership relation $\in$ between the members of that ordinal! Therefore we omit the ordering relation when talking about an ordinal $(X, \leq)$ since it can be derived from $\in$ by the above observation. Another notational convention is that lower case greek letters $\alpha, \beta, \gamma, ...$ are used to denote ordinals.

Examples of ordinals include $\emptyset$, $\{\emptyset\}$ and $\{\emptyset, \{\emptyset\}\}$. We will see later that these are the first three ordinals in some sense and that this enumeration shows the principle of how all ordinals are built up!

2.1 Basic properties of ordinals

The follwing proposition how to get a new ordinal from a given ordinal. The proof is left as an exercise.

PROPOSITION 2.2. *If α is an ordinal, then $\alpha \cup \{\alpha\}$ is an ordinal (with the ordering $x \leq y$ if and only if $[x \in y$ or $x = y]$, as is always the case with ordinals).*

We call the ordinal $\alpha \cup \{\alpha\}$ the *successor* of α and denote it by α^+.

PROPOSITION 2.3. *Each element of an ordinal is an ordinal.*

Proof. Let α be an ordinal and take $x \in \alpha$. So $x = S_\alpha(x)$, so x is a subset of α and therefore well-ordered. Now take any $y \in x$. We prove $S_x(y) = S_\alpha(y) = y$. The second equality follows because α is an ordinal. Take $a \in S_x(y)$. So $a < y$ in the order on $x = S_\alpha(x)$ induced by the order on α. This implies $a \in y = S_\alpha(y)$. Conversely suppose $a \in S_\alpha(y) = y$. So clearly $a < y$ in the well-ordered set α. We have $a \in S_\alpha(x) = x$, because $a < y < x$ in α. So $a, y \in S_\alpha(x) = x$ and since x inherits the ordering from α and $a < y$ in α we have $a < y$ in x, therefore $a < y$ in x. Whence $a \in S_x(y)$.

So we showed that $S_x(y) = y$ for every member of $y \in x$; therefore x is an ordinal. □

PROPOSITION 2.4. *If α is an ordinal, then $\alpha \notin \alpha$.*

Proof. Suppose there is an ordinal α such that $\alpha \in \alpha$. From definition 2.1 we get $\alpha = S_\alpha(\alpha)$. But for each $x \in S_\alpha(\alpha)$ we have $x < \alpha$ and in particular $x \neq \alpha$. With $\alpha \in \alpha = S_\alpha(\alpha)$ we get $\alpha \neq \alpha$, a contradiction. □

The next proposition is crucial in that it shows that the ordering given by the $\in$-relation can be described via (strict) set inclusion.

PROPOSITION 2.5. *Let α, β be ordinals. Then*

$$\alpha \in \beta \Leftrightarrow ([\alpha \subseteq \beta] \text{ and } [\alpha \neq \beta]).$$

Proof. Suppose that $\alpha \in \beta$, then first we have $\alpha = S_\beta(\alpha) \subseteq \beta$ and $\alpha \neq \beta$ by Proposition 2.4. Conversely suppose α is a proper subset of β; moreover we know that α is an ordinal. The set $\beta \setminus \alpha$ is nonempty and therefore has a smallest element γ_0. We claim that $\alpha = \gamma_0$ which implies that $\alpha \in \beta$. Take $\xi \in \gamma_0$. By minimality of γ_0, the element ξ must lie in α.

Conversely take any $\eta \in \alpha$ and suppose that $\eta \notin \gamma_0$. By assumption $\alpha \subseteq \beta$, so $\eta \in \beta$, and $\gamma_0 \in \beta$ by choice of γ_0. Moreover, $\eta \in \alpha$ and $\gamma_0 \notin \alpha$ imply $\eta \neq \gamma_0$. The statements $\xi \notin \gamma_0$ and $\eta \neq \gamma_0$ amount to saying that $\eta \not\leq \gamma_0$ in β. But since β is linearly ordered (even well-ordered), we get $\gamma_0 < \eta$. So $\gamma_0 \in S_\alpha(\eta) \subseteq \alpha$, contradicting the choice of γ_0 being the smallest element of $\beta \setminus \alpha$. □

The main statement about ordinals is that they abide to the **law of trichotomy**:

THEOREM 2.6. *Let α, β be ordinals. Then either $\alpha \in \beta$ or $\beta \in \alpha$ or $\alpha = \beta$, and these cases are mutually exclusive.*

Proof. First, it is very easy to see that Definition 2.1 and Proposition 2.4 imply that the cases are mutually exclusive.

Assume that $\alpha \neq \beta$. We show that

$$\alpha \subseteq \beta \text{ or } \beta \subseteq \alpha \quad (1).$$

Suppose not. Then we have $\alpha \cap \beta \neq \alpha, \beta$. Moreover it is an easy exercise to show that $\alpha \cap \beta$ is an ordinal indeed. So by Proposition 2.5 we get that $\alpha \cap \beta \in \alpha$ and $\alpha \cap \beta \in \beta$, so $\alpha \cap \beta \in \alpha \cap \beta$, contradicting Proposition 2.4. So, (1) is proved.

Finally, (1) together with Proposition 2.5 imply that either $\alpha \in \beta$ or $\beta \in \alpha$. □

This theorem says that we can compare all ordinals by set membership. If $\alpha \in \beta$ holds for two ordinals, the intuition is that α is the smaller and β the bigger ordinal. However, we caution the reader against considering the collection of all ordinals: this is not a set, but a proper class! This result is known as the **Burali-Forti paradox**, and we won't be concerned with it further in this brief note.

For the further development of the theory of ordinals, we need two more results.

PROPOSITION 2.7. *If $\mathcal{A}$ is a set of ordinals, then $\bigcup \mathcal{A}$ is an ordinal.*

Proof. Exercise. $\square$

PROPOSITION 2.8. *Let X be a set of ordinals such that if $\alpha \in X$ and $\nu \in \alpha$ then $\nu \in X$. Then X is itself an ordinal.*

Proof. By proposition 2.7 the set $\beta = \bigcup X$ is an ordinal. Now we distinguish two cases: if each ordinal belonging to X is a member of another ordinal of X then it is easy to check that $X = \beta$. Otherwise, if X contains an ordinal which is not a member of any ordinal belonging to X then that ordinal must be β and it is easy to see that $X = \beta \cup \{\beta\}$. $\square$

2.2 Successor and limit ordinals

When we start with the empty set $\emptyset$ and apply the operation $(-)^+$ again and again, we can build ever larger ordinals. Doing so we see that every ordinal we get this way (except $\emptyset$) is a successor of a (unique) ordinal, its *predecessor*. Those ordinals are said to be *finite ordinals*. We will give a more formal definition of finite ordinals in a moment.

Now, the Axiom of Infinity states that there is a nonempty ordinal λ such that $\alpha \in \lambda$ implies $\alpha^+ \in \lambda$. Such ordinals are called *limit ordinals*. It is easily shown that there is a unique limit ordinal ω in which no element is a limit ordinal (Hint: take a limit ordinal λ and consider the set of limit ordinals in λ). ω is the smallest limit ordinal in the sense that every other limit ordinal contains ω as an element. The members of ω are **by definition** referred to as *finite ordinals*. All other ordinals are said to be *infinite*. Note that "limit" implies "infinite", but not the other way round: $\omega \cup \{\omega\}$, the successor of ω, is infinite, but no limit ordinal.

Indeed, finite ordinals are exactly how the natural numbers are modelled, and ω is a model for $\mathbb{N}$.

3 Ordinals are woset prototypes

The aim of this section is to prove that every well-ordered set is order-isomorphic to a unique ordinal. We deal with the uniqueness part first.

PROPOSITION 3.1. *If α and β are ordinals and $f : \alpha \to \beta$ is an order isomorphism, then $\alpha = \beta$ and f is the identity map.*

Proof. If $\alpha \neq \beta$ we may assume without loss of generality that $\alpha \in \beta$ by theorem 2.6. But since α is also a segment of β, we get that the woset β is isomorphic to one of its segments, contradicting Corollary 1.2. So $\alpha = \beta$. Now suppose that f is not the identity. By 1.1 we have $x \leq f(x)$. So there is $x \in \alpha = \beta$ with $x < f(x)$. Let x_0 be the least such. Then we claim that $x_0 \notin \mathrm{im}(f)$: For all $y < x_0$ we have $f(y) = y < x_0$, and for $z \geq x_0$ we get $f(z) \geq f(x_0) > x_0$. So f is not surjective, whence no order-isomorphism. $\square$

The strategy for proving that every woset is order-isomorphic to some ordinal is the following. First we show that if for each segment there is an ordinal isomorphic to that segment, then so is the woset. This will be a bit technical. Then the easier step is to show that for every woset, for each segment there **is** indeed an ordinal isomorphic to that segment. Interestingly, this second result uses the first, and the two of them combined will give the desired result that every woset is isomorphic to an ordinal. This strategy could well be called “woset induction”.

PROPOSITION 3.2. *If every segment of a woset $(X, \leq)$ is isomorphic to an ordinal, then $(X, \leq)$ is itself isomorphic to an ordinal.*

Proof. The statement of the proposition is clearly valid if $X = \emptyset$ so we assume $X \neq \emptyset$. For each $a \in X$ let $C(a)$ be the unique ordinal isomorphic to it and let $\varphi_a : S_X(a) \to C(a)$ be the unique order isomorphism from $S_X(a)$ to $C(a)$. (To see that there is only one order isomorphism from $S_X(a)$ to $C(a)$ suppose that there are two different ones, η and ψ. So then $eta \circ \psi^{-1}$ is an order isomorphism from $C(a)$ to itself which is not the identity, contradicting Proposition 3.1).

We split the remainder of the proof into several subclaims.

Claim 1. If $a < b$ in X, then $\varphi_b(a) = C(a)$.

We have to be aware of the double role of the elements of ordinals: they are segments as well! Restricting $\varphi_b(a)$ to $S_{S_X(b)}(a) = S_X(a)$ gives an isomorphism between $S_X(a)$ and the segment of $\varphi_b(a)$ - which is equal to $\varphi_b(a)$ since $C(b)$ is an ordinal! Let us denote that isomorphism with ψ. So

$\varphi_a \circ \psi^{-1} : \varphi_b(a) \to C(a)$ is an order isomorphism which implies $\varphi_b(a) = C(a)$ by Proposition 3.1. So Claim 1 is proved.

Let $W = \{C(a) : a \in X\}$.

Claim 2. W is an ordinal.

By Proposition 2.8 it suffices to prove the following: if $\beta \in W$ and $\alpha \in \beta$ then $\alpha \in W$. There is $y \in X$ such that $\varphi(y) : S_X(y) \to C(y) = \beta$ is an order isomorphism. Since α in β there is $x \in S_X(y)$ such that $\alpha = \varphi_y(x) = C(x)$ by Claim 1. So Claim 2 is proved.

Claim 3. The map $x \mapsto C(x)$ establishes an order-isomorphism between X and W.

It suffices to show that $x < y$ in X entails $C(x) \in C(y)$ in W. (Surjectivity of C is given by definition of W). If $x < y$ then $x \in S_X(y)$ so $C(x) = \varphi_y(x) \in C(y)$. This establishes Claim 3 and the proof of the theorem. □

PROPOSITION 3.3. *In a woset, every segment is isomorphic to some ordinal.*

Proof. If this is not true for a particular woset $(X, \leq)$, then assume that $y \in X$ is the smallest element for which $S_X(y)$ is not isomorphic to any ordinal. Note that $S_X(y)$ is a woset in its own right. So for all $z \in S_X(y)$ the segment $S_{S_X(y)}(z) = S_X(z)$ is isomorphic to some ordinal by minimality of y. But then, by Propositon 3.2, $S_X(y)$ is isomorphic to an ordinal, contradicting our assumption. □

Combining Propositions 3.1, 3.2 and 3.3 we get

THEOREM 3.4. *Every well-ordered set is order-isomorphic to a unique ordinal.*

The unique ordinal to which a given woset $(X, \leq)$ is isomorphic is usually denoted by $\mathrm{Ord}(X, \leq)$.

4 The Axiom of Choice

The Axiom of Choice (AC) is the most argued about axiom of all axioms of set theory. This section aims at giving an introduction to AC and deriving two important consequences from it: Zorn's Lemma and the Well-Ordering Theorem (both of which are in fact equivalent to AC, but we will not show this).

Axiom of Choice (AC). *If C is a set whose members are pairwise disjoint nonempty sets, then there is a set $s \subseteq \bigcup C$ (the "choice set") such that*

$$(\forall c \in C)(\exists x \in c) : s \cap c = \{x\}.$$

Intuitively, AC is stating that for any collection of non-empty sets we can pick exactly one element of each member set. The point is that AC doesn't give us a recipe how to do so. That's why it is called a non-constructive axiom; and indeed it is not derivable from the other axioms of set theory.

In the following we give another formulation of AC which is more commonly used.

Axiom of Choice, functional version. *Let C be a nonempty set consisting of nonempty sets. Suppose I is a set and $z : I \to C$ is any function. Then there is a* choice function *$f : I \to \bigcup C$ such that $f(i) \in z(i)$ for all $i \in I$.*

PROPOSITION 4.1. *The Axiom of Choice is equivalent to its functional version.*

Proof. $\Longrightarrow$: Consider the set $D = \{z(i) \times \{i\} : i \in I\}$. It is easy to see that the members of D are pairwise disjoint. So there is a choice set $s \subseteq \bigcup D$ such that for all $d \in D$ there is $x \in d$ such that $s \cap d = \{x\}$. Notice that x has the form (y, i) for some $y \in z(i)$. So we define $f : I \to \bigcup C$ in the following way:

$$f = \{(i, y) : c \in C \text{ and } (y, i) \in s \cap (z(i) \times \{i\})\}.$$

Since $s \cap (z(i) \times \{i\})$ contains only one element, the relation defined above is indeed a function, and it has the desired property.

$\Longleftarrow$: Let C be a collection of pairwise disjoint sets. In the notation of above choose $I := C$ and $z := id$. Then there is $f : C \to \bigcup C$ such that $f(c) \in c$ for all $c \in C$. Then $im(f) = f(C)$ is the desired choice set. $\square$

The functional version of (AC) can also be formulated in a less formal way:

> If I is a set and $\{C_i : i \in I\}$ is a family of nonempty sets, then there is a choice function $f : I \to \bigcup_{i \in I} C_i$ such that $f(i) \in C_i$ for all $i \in I$.

The family of all such choice functions is commonly denoted by $\prod_{i \in I} C_i$ and called *product* of the C_i. A very short version of (AC) thus states that products of nonempty sets are nonempty.

Zorn's Lemma (ZL) deals with maximal elements in posets. If $(P, \leq)$ is a poset then $a \in P$ is called *maximal* if $b \geq a$ implies $b = a$. A *chain* in P is a subset $C \subseteq P$ that is totally ordered with respect to the induced ordering. If $A \subseteq P$ we call $x \in P$ an *upper bound* of A if for all $a \in A$ we have $a \leq x$.

Zorn's Lemma (ZL). *Let $(P, \leq)$ be a non-empty poset and suppose that every chain in P has an upper bound in P. Then P has a maximal element.*

THEOREM 4.2. (AC) *implies* (ZL).

Proof. Let $(P, \leq)$ be a non-empty poset satisfying the conditions of (ZL). Let $\mathcal{C}$ denote the set of all chains in P. For any chain C in P let

$$B(C) = \{t \in P \setminus C : t \text{ is an upper bound of } C\}$$

. If some $B(C)$ is empty, then we know that C has an upper bound and this must be the greatest element of C which is necessarily maximal in P, so we are done.

So we consider the case that $B(C)$ is nonempty for all C and lead this case to a contradiction. By the functional version of (AC), there is a function $\psi : \mathcal{C} \to P$ such that $\psi(C) \in B(C)$ for all chains $C \in \mathcal{C}$. In other words, for every chain C, the element $\psi(C) \in P$ is an upper bound of C and

$$\psi(C) \notin C \quad [\star].$$

Now, let $\mathcal{W}$ denote the set of subsets $W \subseteq P$ with the properties:

(a) W is a well-ordered chain in P.

(b) For every segment $S_W(x) \in P$ we have that $w = \psi(S_W(x))$

Claim. $T := \bigcup \mathcal{W}$ is itself well-ordered.

We do this by showing that if V, W are members of W then one is a segment of the other and then concluding that $\bigcup \mathcal{W}$ is well-ordered by Proposition 1.3.

So let $V, W \in \mathcal{W}$ be distinct. We may assume $W \not\subseteq V$. Let w_0 be the smallest element of $W \setminus V$. We claim that $S_W(w_0) = V$. Clearly $\subseteq$ holds because of minimality of w_0. On the other hand suppose that equality is false. Let v_0 be the smallest element of $V \setminus S_W(w_0)$. Clearly $S_V(v_0) \subseteq S_W(w_0)$ by minimality of v_0. If $S_V(v_0) = S_W(w_0)$ then by definition of ψ and $\mathcal{W}$ we have $w_0 = \psi(S_W(w_0)) = \psi(S_V(v_0)) = v_0$, contradicting $w_0 \notin V$. But if $S_V(v_0)$ is a proper subset of $S_W(w_0)$ then there is w_1 minimal in $S_W(w_0)$ with $w_1 \notin S_V(v_0)$. So $S_{S_W(w_0)}(w_1) = S_W(w_1) = S_V(v_0)$ and again by definition of $\mathcal{W}$ we get $w_1 = \psi(S_W(w_1)) = \psi(S_V(v_0)) = v_0$, contradicting $v_0 \notin S_W(w_0)$. In any way we get a contradiction, so the assumption that $S_W(w_0)$ is a proper subset of V is false. Therefore $V = S_W(w_0)$ and the claim is proved.

Now we are almost done: T is now itself a well-ordered chain, and so is $U := T \cup \{\psi(T)\}$. It is indeed easy to check that $T \cup \{\psi(T)\}$ also satisfies condition (b). So $U \in \mathcal{W}$ which implies that $\psi(T) \in U \subseteq \bigcup \mathcal{W} = T$ contradicting [$\star$] which followed from assuming that P has no maximal elements. So P has maximal elements and (ZL) is proved. $\square$

We next look at the Well-Ordering Theorem - which is actually an axiom; but sometimes the statement of this axiom is confused with the *theorem* stating that (ZL) implies the Well-Ordering Theorem.

Well-Ordering Theorem (WO). *Every set can be well-ordered.*

THEOREM 4.3. (ZL) *implies* (WO).

Proof. Let X be any set and let R be any relation on X. We define the *domain* of r to be

$$\mathrm{dom}(r) = \{x \in X : (\exists y \in X) : xry\}$$

and analogously the *range* of r to be $\mathrm{ran}(r) = \{y \in X : (\exists x \in X) : xry\}$. A *partial well-ordering* on X is a relation r on X such that $\mathrm{dom}(r) = \mathrm{ran}(r)$ and r is a well-ordering on $\mathrm{dom}(r) = \mathrm{ran}(r)$. We denote the set of all partial well-orderings on X by PWO. For two partial well-orderings r_1, r_2 we define

$$
\begin{array}{ll}
r_1 \preceq r_2 \text{ iff } & r_1 = r_2 \text{ or} \\
& [\ \mathrm{dom}(r_1) = S_{\mathrm{dom}(r_2)}(y^*), \text{ where } y^* \text{ denotes the} \\
& \text{smallest member of } \mathrm{dom}(r_2) \setminus \mathrm{dom}(r_1), \\
& \text{and} \\
& r_1 = r_2 \cap (\mathrm{dom}(r_1) \times \mathrm{dom}(r_1))\ ].
\end{array}
$$

It is easy to verify that $(PWO, \preceq)$ is a poset satisfying the conditions of (ZL). So there is a maximal partial well-ordering m. Again, it is easy to check that maximality of m implies $\mathrm{dom}(m) = X$. So m well-orders X and we are done. □

Finally we want to show that (WO) in turn implies (AC).

THEOREM 4.4. (WO) *implies* (AC).

Proof. Let C be a collection of pairwise disjoint nonempty sets. Using (WO), well-order $\bigcup C$ and call the resulting relation $\leq$. Now define the "choice set" $s \subseteq \bigcup C$ by

$$s = \{x \in \bigcup C : (\exists m \in C) : x \text{ is the smallest element of } m\}.$$

Clearly, s "picks" exactly one element of every member m of C. □

Note that in the above proof, the well-ordering provided us with a **method** to choose exactly one element from each member of C, namely: "pick the smallest member of each set in S". Such a choice can not be made if we just assume the axioms of ZF without the Axiom of Choice (or one of its equivalents, of which we've seen (ZL) and (WO)). Bertrand Russell once brought forth the following analogy for the Axiom of Choice: You don't need the Axiom of Choice if you want to pick exactly one shoe from each pair of an arbitrarily large collection of pairs of shoes: from every pair of shoes, take the left shoe. But you need the Axiom of Choice when you want to do the same with socks (assuming that there is no notion of "left sock" and "right sock"). Then you have no rule guiding you through a choice so to speak.

5 The first uncountable ordinal ω_1

The goal of this chapter is to show that there is some ordinal λ which is *uncountable*, that is, it is infinite but there is no bijection from ω onto λ.

The construction of λ requires (AC). The strategy is to show that there actually is an uncountable set X, then we well-order this set using (WO) yielding a woset $(X, \leq)$. We then consider the ordinal $\lambda := \mathrm{Ord}(X, \leq)$ using Theorem 3.4 and conclude that λ is uncountable. In order to construct X we need a theorem that is due to Georg Cantor, the founder of set theory. The proof is an absolute marvel in mathematics; it has been elected to be one of the ten most beautiful proofs ever conceived.

Cantor's Theorem. *For any set X let $\mathcal{P}(X)$ denote its power set, that is the set of all subsets of X. Then for no set X there is a surjection $s : X \to \mathcal{P}(X)$.*

Proof. Suppose there is such a surjection s. Consider the following subset of X:

$$M := \{t \in X : t \notin s(t)\}$$

Since $M \in \mathcal{P}(X)$ and s is surjective there is some t_0 such that $s(t_0) = M$. But then if $t_0 \in M$ we get $t_0 \notin M$ by the very definition of M. And if $t_0 \notin M$ we get $t_0 \in M$ again by the very definition of M. So the assumption that s is surjective leads to a contradiction. $\square$

As outlined in the beginning of the Appendix, we are now able to state that $\mathcal{P}(\omega)$ is uncountable. With Theorem 4.3 we can endow it with a well-ordering $\leq$ and thus can consider the ordinal

$$\lambda := \mathrm{Ord}(\mathcal{P}(\omega))$$

which must be uncountable. Now distinguish two cases:

- λ contains but finite and countable ordinals. Then set $\omega_1 = \lambda$.
- λ contains uncountable ordinals. Then let ω_1 be the smallest such.