Well – ordered sets and Zorn's Lemma

In the remainder of this course we shall focus our attention on two topics:

1. The arithmetic of transfinite cardinal numbers and the tools needed for this.
2. Set – theoretic assumptions that are equivalent to the Axiom of Choice.

We begin with a simple question about the ordering of transfinite cardinal numbers: ***Given a set*** S***, is the set of cardinal numbers*** $\mathrm{Card}(S)$ ***linearly ordered?*** Common sense suggests that it should be, but we have seen that even proving the antisymmetric property for cardinal numbers required a nontrivial result (the Schröder – Bernstein Theorem). It turns out that proving the ordering is linear requires even more sophisticated input.

<u>Well – Ordering Axiom.</u> If X is a nonempty set, then there exists a partial ordering relation $\mathcal{O}$ on X such that X is well – ordered with respect to $\mathcal{O}$.

When working with a given well – ordering on examples, we shall use the standard symbolism $x < y$ and $x \leq y$ in our discussions.

One unsettling aspect of this assumption is that ***no one has ever constructed an explicit well – ordering of the real number system***. However, the Well – Ordering Axiom is logically equivalent to the Axiom of Choice. A proof that the latter implies the existence of a well – ordering is presented in Sections 7.1 and 7.3 of Cunningham. On the other hand, deriving the Axiom of Choice from the Well – Ordering Axiom is a fairly straightforward exercise.

<u>PROOF THAT WELL – ORDERING IMPLIES THE AXIOM OF CHOICE.</u> Given a nonempty set A, take some well – ordering of this set. If B is a nonempty subset of A, then by well – ordering it has a least element, and we can simply define $c(B) \in B$ to be this least element.■

<u>Countable well – ordered sets.</u> The most basic examples of well – orderings on a countable set are the usual well – ordering of $\mathbb{N}$ and of its subsets $\{1, 2, \ldots, n\}$, but there are also many others. Section 8.3 of Cunningham describes arithmetic constructions on well – ordered sets which yield many additional examples (see also Figure 8.1 in the next section of the text).

The following properties of well – ordered sets will play important roles in our uses of such objects.

<u>Theorem (Well – ordering of subsets).</u> *A nonempty subset of a well – ordered set is well – ordered with respect to the induced ordering. Furthermore, if X is a well – ordered set and $Y \subset X$ is a subset such that $|Y| < |X|$, then the induced ordering on Y is in $1-1$ order – preserving correspondence with a subset of the form $\{y \in X \mid y < x_0\}$ for some $x_0 \in X$.*

<u>PROOF.</u> The first sentence is easy to verify, for if W is a nonempty subset of Y then it is also a nonempty subset of X and therefore W has a least element.

The proof of the second part of the theorem will be shown by a recursive procedure known as ***<u>transfinite recursion</u>***. As in the finite case, the basic idea is that we have a construction which has been completed for all $b < a$ in a well – ordered set X, and using this construction we describe a way to find a decent extension to the case $b = a$. If a is not a maximal element of A, then the set of elements which are strictly greater than a has a minimal element which we shall call $\sigma(a)$ or $a+1$. Then one uses a similar procedure to extend the construction to a, and so on.

Suppose now that Y is a nonempty subset of a well – ordered set X and $|Y| < |X|$. For each $a \in X$ we want to construct a $1-1$ onto, order – preserving, map h_a from $\{y \in X \mid y < a\}$ into Y with the following properties:

1. If z_Y and z_X are the least elements of Y and X respectively, then h_a sends z_X to z_Y.
2. If $b < a$, then h_b is the restriction of h_a to $\{y \in X \mid y < b\}$.
3. If h_b is defined for all $b < a$ and the union of the images of these maps is a proper subset of X, then $h_b(a)$ is the first element of Y not in the images of the maps h_b.

The first step shows how one begins the recursive process, the second one describes what one wants at each step, and the third indicates what the next step should be or if the recursive process must be terminated. Each step is fairly straightforward, but eventually it becomes impossible to complete the third step, and this must happen by the time that that the cardinality of $\{y \in X \mid y < b\}$ becomes greater than $|Y|$. As in the finite case, the existence of the desired mapping is given as follows:

PRINCIPLE OF TRANSFINITE INDUCTION. *Let W be a nonempty well – ordered set, and suppose that we are given a sequence of statements S_α where α runs through all the elements of W. Furthermore, assume also that the following hold:*

1. *Statement S_z is true where z denotes the least element of W.*
2. *For all $\alpha \in W$, if Statement S_β is true for* **ALL** *$\beta < \alpha$ then Statement S_α is also true.*

Then all of the statements S_α are true.

The justification is similar to the corresponding argument for the Strong Principle of Finite Induction.■

Returning to the proof of the theorem, if c is the first element of X such that h_c cannot be defined, then the union of the h_a for $a < c$ will yield the desired $1-1$ correspondence between Y and the set of all $a \in X$ such that $a < c$.■

The preceding theorem implies that the set $\mathbf{Card}(S)$ is linearly ordered.

Theorem (Linear ordering of cardinal numbers). *If X and Y are subsets of some set S, then either $|X| \leq |Y|$ or $|Y| \leq |X|$.*

PROOF. Let W be a set such that $X \cup Y \subset W$ and $|W| > |X \cup Y|$; for example, we can take $W = \mathcal{P}(X \cup Y)$. Choose a well – ordering of W. By the previoius theorem we know that X and Y are in $1-1$ correspondences with subsets of the form $\{w \in W \mid w < a\}$ and $\{w \in W \mid w < b\}$ respectively for some a and b in W. If $a = b$ then we automatically have $|X| \leq |Y|$. Otherwise, if $a \neq b$, then every well – ordering is a linear ordering because the set $\{a, b\}$ has a least element, so either a or b is a minimal element; if a is minimal then $|X| \leq |Y|$, while if b is minimal then $|Y| \leq |X|$.■

Zorn's Lemma

We shall now consider a statement which at first might look bizarre, but it is actually equivalent to the Axiom of Choice and the Well – Ordering Axiom. The statement turns out to be extremely useful for giving (non – constructive) proofs that certain sorts of objects and functions exist.

ZORN'S LEMMA. Let X be a nonempty partially ordered set such that each linearly ordered subset Y has an upper bound in X. Then X has a maximal element.

Section 7.1 of Cunningham indicates how one can derive this statement from the Axiom of Choice. Since we have seen that well – ordering implies the Axiom of Choice, it follows that the Well – Ordering Axiom also implies Zorn's Lemma. We shall explain how Zorn's Lemma implies the Axiom of Choice and the Well – Ordering Axiom in an addendum to these notes, and we shall try to motivate Zorn's Lemma by giving a quick and dirty proof here:

<u>DERIVING OF ZORN'S LEMMA FROM THE AXIOM OF CHOICE AND THE WELL – ORDERING AXIOM.</u> Suppose that X is a partially ordered set satisfying the hypothesis of Zorn's Lemma, and consider what would happen if the conclusion were false. <u>*In the argument that follows it will be convenient to extend the notation for closed, open and half open intervals from the real numbers to an arbitrary linearly ordered set.*</u>

As suggested in the previous theorem, take a well – ordering of the power set $\mathcal{P}(X)$. We claim it is possible to define a strictly increasing map f from $\mathcal{P}(X)$ to X by transfinite recursion. If we can do this, we will have a contradiction because there is no $1-1$ map from $\mathcal{P}(X)$ to X. Let $k : X \rightarrow \mathcal{P}(X)$ be a choice function.

If z_X is the least element of X, define $f(z_X) = k(X)$ to begin the process. Suppose now that we have defined the function on $[0, \alpha)$, and let $J_\alpha = f[[0, \alpha)]$. By hypothesis the latter is a linearly ordered subset of X and as such it has an upper bound. Use the choice function k to select a particular upper bound $u(\alpha)$. We are also assuming that X has no maximal element so the set of all elements strictly greater than $u(\alpha)$ is nonempty; use k again to select some $f(\alpha) > u(\alpha)$. Since f is strictly increasing for $\beta < \alpha$ and $f(\alpha)$ is greater than every element of J_α by construction, it follows that f is $1-1$ on the closed interval $[0, \alpha]$. This completes the recursive step in the definition of the strictly increasing map $f : \mathcal{P}(X) \longrightarrow X$.

As noted in the second paragraph of the argument, this yields a contradiction. Where is the problem? The construction of f relies heavily on the fact that X has no maximal element, so this must be false. Thus X must have a maximal element, and the existence of such an element is exactly what is needed to prove Zorn's Lemma.■

In the next lecture we shall indicate how Zorn's Lemma has major implications for the arithmetic of cardinal numbers, but first we shall illustrate the use of this result to extend a result about finite partial orderings to the transfinite case.

<u>Theorem (Extending partial orderings).</u> *Let A be a set, and let $\mathcal{O} \subset A \times A$ be a partial ordering. Then there is a linear ordering $\mathcal{L} \subset A \times A$ such that $\mathcal{O} \subset \mathcal{L}$.*

We frequently say that $\mathcal{L}$ is a ***compatible linear ordering*** or $\mathcal{L}$ *is compatible with* $\mathcal{O}$.

As noted above, a result of this type is useful for many purposes. For example, if X is a finite set and A is a family of subsets of X, then sometimes one wants prove a fact about the elements of A by mathematical induction, where A is linearly ordered such that for each pair of elements B, C in A such that $B \subset C$ we also have $B < C$.

The nonconstructive nature of the stated theorem is illustrated by one simple fact: ***A compatible linear ordering for the set*** $\mathcal{P}(\mathbb{N})$ ***of subsets of the natural numbers (ordered by inclusion) has never been explicitly constructed.*** In contrast, given an arbitrary partial ordering $\mathcal{O}$ on a finite set, we have algorithmically constructed some explicit linear orderings $\mathcal{L}$ which are compatible with $\mathcal{O}$.

PROOF. We shall first show that there is a maximal partial ordering containing the given one and then show that such a maximal partial ordering must be a linear ordering.

Let $\mathbf{C}$ be the collection of all partial orderings of A that contain $\mathcal{O}$. Then $\mathbf{C}$ is partially ordered by set − theoretic inclusion. Let $\mathbf{D}$ be a subset of $\mathbf{C}$ that is linearly ordered by inclusion. If we can show that $\mathbf{D}$ has an upper bound in $\mathbf{C}$, then Zorn's Lemma will imply that $\mathbf{C}$ has a maximal element.

Denote the elements of $\mathbf{D}$ by $\mathcal{L}_s$ where s runs through some indexing set S, and let $\mathcal{L}$ be the union of all the sets $\mathcal{L}_s$. Clearly $\mathcal{L}$ contains $\mathcal{O}$ since each $\mathcal{L}_s$ does; we would like to show that $\mathcal{L}$ is also a partial ordering. The relation $\mathcal{L}$ is reflexive because $\mathcal{L}$ contains $\mathcal{O}$ and $\mathcal{O}$ is reflexive.

To verify the relation $\mathcal{L}$ is antisymmetric, suppose that both (a, b) and (b, a) belong to $\mathcal{L}$. Then there are partial orderings $\mathcal{L}_s$ and $\mathcal{L}_t$ such that (a, b) belongs to $\mathcal{L}_s$ and (b, a) belongs to $\mathcal{L}_t$. Since $\mathbf{D}$ is linearly ordered by inclusion it follows that one of $\mathcal{L}_s$ and $\mathcal{L}_t$ contains the other. If $\mathcal{L}_u$ is the larger relation, then both (a, b) and (b, a) belong to $\mathcal{L}_u$, and since the latter is a partial ordering this means that $a = b$.

Finally, suppose that both (a, b) and (b, c) belong to $\mathcal{L}$. Then there are partial orderings $\mathcal{L}_s$ and $\mathcal{L}_t$ such that (a, b) belongs to $\mathcal{L}_s$ and (b, c) belongs to $\mathcal{L}_t$. Since $\mathbf{D}$ is linearly ordered by inclusion it follows that one of $\mathcal{L}_s$ and $\mathcal{L}_t$ contains the other. If $\mathcal{L}_u$ is the larger relation, then both then both (a, b) and (b, c) belong to $\mathcal{L}_u$, and since the latter is a partial ordering this means that (a, c) belongs to $\mathcal{L}_u$, which is contained in $\mathcal{L}$. Therefore $\mathcal{L}$ is a partial ordering. By construction, it is an upper bound for the elements of $\mathbf{D}$, and thus Zorn's lemma implies that $\mathbf{C}$ must have a maximal element.■

The second part of the proof of the theorem is contained in the following result:

Proposition (Maximal partial orderings are linear orderings). *Let A be a set, and let $\mathcal{O} \subset A \times A$ be a maximal partial ordering. Then $\mathcal{O}$ is a linear ordering.*

PROOF. Suppose that $\mathcal{O}$ is not a linear ordering. Then we can find x, y in A such that neither (x, y) nor (y, x) lies in $\mathcal{O}$. We shall obtain a contradiction by expanding $\mathcal{O}$ to a partial ordering that contains (x, y). In order to express the argument in familiar notation we shall write $u \le_{\mathcal{O}} v$ to signify that (u, v) lies in $\mathcal{O}$.

Define a new binary relation $\mathcal{L}$ such that (u, v) lies in $\mathcal{L}$ if and only if either $u \le_{\mathcal{O}} v$ or else both $u \le_{\mathcal{O}} x$ and $y \le_{\mathcal{O}} v$. The proof of the proposition then reduces to showing that $\mathcal{L}$ is a partial ordering.

The relation $\mathcal{L}$ is reflexive. Since $\mathcal{O}$ is a partial ordering, for each $a \in A$ we know that $(a, a) \in \mathcal{O} \subset \mathcal{L}$.

The relation $\mathcal{L}$ is transitive. Suppose that $(a, b) \in \mathcal{L}$ and $(b, c) \in \mathcal{L}$. Then there are two options for each of the ordered pairs in the preceding sentence and thus a total of four separate cases to consider:

1. We have $a \le_{\mathcal{O}} b$ together with $b \le_{\mathcal{O}} c$.
2. We have $a \le_{\mathcal{O}} b$ together with both $b \le_{\mathcal{O}} x$ and $y \le_{\mathcal{O}} c$.
3. We have both $a \le_{\mathcal{O}} x$ and $y \le_{\mathcal{O}} b$ together with $b \le_{\mathcal{O}} c$.
4. We have both $a \le_{\mathcal{O}} x$ and $y \le_{\mathcal{O}} b$ together with both $b \le_{\mathcal{O}} x$ and $y \le_{\mathcal{O}} c$.

In the first case, since $\mathcal{O}$ is a partial ordering we have $a \le_{\mathcal{O}} c$, so that $(a, c) \in \mathcal{O}$. In the second case, since $\mathcal{O}$ is a partial ordering we have $a \le_{\mathcal{O}} x$, and therefore (a, c) satisfies the second criterion to be an element of $\mathcal{L}$. In the third case, since $\mathcal{O}$ is a partial ordering we have $y \le_{\mathcal{O}} c$, and therefore (a, c) satisfies the second criterion to be an element of $\mathcal{L}$. Finally, in the fourth case since $\mathcal{O}$ is a partial ordering the middle two conditions imply that $y \le_{\mathcal{O}} x$, which contradicts our original hypothesis that neither of the relations $x \le_{\mathcal{O}} y$ or $y \le_{\mathcal{O}} x$ is valid. Therefore the fourth case is impossible, and this completes the proof of transitivity.

The relation $\mathcal{L}$ is antisymmetric. Suppose that $(a, b) \in \mathcal{L}$ and $(b, a) \in \mathcal{L}$. Then we have the same four cases as in the proof of transitivity, the only difference being that one must replace c by a in each case. In the first case, since $\mathcal{O}$ is a partial ordering we must have $a = b$. In all the remaining cases, since $\mathcal{O}$ is a partial ordering the given conditions combine to imply $y \le_{\mathcal{O}} x$, which contradicts the assumption on $\mathcal{L}$. Thus only the first case is possible, and this completes the proof that the relation $\mathcal{L}$ is antisymmetric.■

A more general perspective. For many decades mathematicians have found Zorn's Lemma to be particularly effective for proving theorems that depend upon the Axiom of Choice, largely because most of these results translate easily into the existence of a maximal object of some sort. From this perspective, the proofs usually have two distinct parts:

1. Showing that a maximal object of some type must exist using Zorn's Lemma.
2. Showing that such maximal objects must have certain desired properties.

In the interests of illustrating proofs using Zorn's Lemma, here is one more example:

Theorem (Hausdorff Maximal Principle). *Every nonempty partially ordered set contains a maximal linearly ordered subset.*

PROOF. Let X be a nonempty partially ordered set, let $\mathcal{R}$ be the partial ordering and consider the family Y of all subsets A of X such that

$$\mathcal{R}|A \quad = \quad \mathcal{R} \cap A \times A$$

Is a linear ordering on A, with the partial ordering of Y given by set $-$ theoretic inclusion. The family Y is nonempty, for if $x \in X$ then one has the trivial linear ordering

$$\{x\} \times \{x\} \quad = \quad \mathcal{R} \cap (\{x\} \times \{x\})$$

on the one point subset $\{x\} \subset X$.

Suppose that we have a linearly ordered subfamily of subsets X_a as above. If we take $W = \cup_a X_a$ then we claim that $\mathcal{T} = \mathcal{R}|W$ is a linear ordering on W. By construction it is a partial ordering, so the only thing to prove is the dichotomy property. Suppose now that $x, y \in W$. Then one can find a and b such that $x \in X_a$ and $y \in X_b$. The linear ordering property implies that one of a or b is greater than or equal to the other; if c denotes this element, then we have $x, y \in X_c$. Since the latter set is linearly ordered with respect to

$$\mathcal{S}_c \quad = \quad \mathcal{R}|X_c$$

it follows that either $(x, y) \in \mathcal{S}_c$ or $(y, x) \in \mathcal{S}_c$, and since the latter is contained in $\mathcal{T}$ it follows that one of the two pairs must lie in $\mathcal{T}$. Therefore $\mathcal{T}$ is a linear ordering, and therefore W is an upper bound in Y for all of the linearly ordered subsets X_a.

We can now use Zorn's Lemma to conclude that Y has a maximal element, which is given by a subset M with the linear ordering $\mathcal{L} = \mathcal{R}|M$. It follows immediately that M is a maximal linearly ordered subset.■

As in the preceding theorem, if X is finite there are much easier ways to find examples of maxial linearly ordered subsets. One method is to take a linearly ordered subset with the greatest number of elements.

For the sake of completeness we note that ***the Hausdorff Maximal Principle is also logically equivalent to Zorn's Lemma*** (or the Axiom of Choice or the Well – Ordering Principle). However, the property of extending partial orderings to linear orderings turns out to be logically weaker than the Axiom of Choice and its equivalent statements (this result is due to A. R. D. Mathias).

.