Mathematics 6310
Zorn's lemma
Ken Brown, Cornell University, September 2010

1. Statement and first proof

Let X be a poset (partially ordered set). A *chain* in X is a totally ordered subset, i.e., a subset in which any two elements are comparable. An element $m \in X$ is called *maximal* if there is no $x \in X$ with $x > m$. Note that a maximal element is not necessarily a largest element, which would be an element m such that $x \leq m$ for all $x \in X$. *Zorn's lemma* is the following result:

Theorem 1. *Let X be a poset in which every chain has an upper bound. Then X has at least one maximal element.*

There is a very short, straightforward proof of Zorn's lemma that uses ordinal numbers. The ordinal numbers extend the natural numbers:

$$0, 1, 2, \ldots, \omega, \omega + 1, \omega + 2, \ldots, \omega + \omega =: \omega 2, \omega 2 + 1, \ldots, \omega 3, \ldots, \omega^2, \ldots.$$

They go on forever. More precisely, given any set of ordinals, there is an ordinal bigger than all of them. In particular, there is no such thing as the set of *all* ordinals. The ordinals are well-ordered: Any nonempty collection of ordinals has a smallest element. This justifies proofs and definitions by induction, just as for the natural numbers. See the separate handout The ordinal numbers for a more detailed outline of the theory of ordinals.

To prove Zorn's lemma, it will be convenient to assume that we have a a "successor operation" on X, denoted $x \mapsto x^+$, such that $x^+ > x$ if x is not maximal, and $x^+ = x$ if x is maximal. (The axiom of choice guarantees that there is indeed such a function.) It will also be convenient to assume that we have a function that associates a "canonical" upper bound to any chain. (Again, such a function always exists by the axiom of choice. In many applications one has least upper bounds and can use these instead of appealing to the axiom of choice.)

Now let X be as in the statement of Zorn's lemma. We define inductively a weakly increasing sequence (x_α) in X, indexed by the ordinals. Suppose α is an ordinal such that x_β has already been defined for $\beta < \alpha$. If α has an immediate predecessor β (i.e., $\alpha = \beta + 1$), we take x_α to be the successor x_β^+. Otherwise, we take x_α to be the canonical upper bound of the chain $(x_\beta)_{\beta<\alpha}$.

Note that the sequence x_α is *strictly* increasing if and only if X has no maximal element. But it can't be strictly increasing, because then the ordinals would be in 1–1 correspondence with a subset of X, contradicting the fact that one cannot form the set of all ordinals. This proves that X has a maximal element and completes the proof of Zorn's lemma.

The rest of this handout will describe an alternative proof of Zorn's lemma that doesn't use ordinals but is longer and somewhat less intuitive. We begin with a reformulation of Theorem 1.

2. The Hausdorff maximal principle

The Hausdorff maximal principle is the following result:

Theorem 2. *Every poset contains a maximal chain (i.e., a chain that is not contained in any bigger chain).*

This is an easy consequence of Zorn's lemma. Indeed, let X be an arbitrary poset and let $\mathcal{X}$ be the set of all chains in X, ordered by inclusion. Then $\mathcal{X}$ satisfies the hypothesis of Zorn's lemma because if $\mathcal{C} \subseteq \mathcal{X}$ is a chain in $\mathcal{X}$, then $\bigcup_{C \in \mathcal{C}} C$ is easily seen to be a chain in X and hence an upper bound for $\mathcal{C}$ in $\mathcal{X}$.

Conversely, one can easily deduce Zorn's lemma from Theorem 2: If X is as in Theorem 1, let C be a maximal chain. Then C has an upper bound $m \in X$, and maximality implies that $m \in C$ and hence is the largest element of C. Another application of the maximality of C now implies that m is a maximal element of X.

So we can prove either Theorem 1 or Theorem 2, whichever we choose.

3. Strategy of the proof

The proof of Zorn's lemma that I will give is adapted from the proofs in Lang, *Real and Functional Analysis*, and Halmos, *Naïve Set Theory*. The idea is to surreptitiously construct the set $M = \{x_\alpha\}$ (notation as in Section 1), without ever mentioning ordinals. Thus instead of building M step by step, we will give an abstract description of it. Consider subsets $N \subseteq X$ with the following closure properties:

(i) If $x \in N$, then $x^+ \in N$.
(ii) For any chain $C \subseteq N$, the canonical upper bound of C is in N.

For brevity, call N *closed* if it satisfies (i) and (ii). Note that X itself is closed, for example, but the empty set is not closed. [It doesn't satisfy (ii).] Note also that the intersection of any family of closed sets is closed. In particular, the intersection of *all* closed sets is closed. Call it M; it is then the smallest closed set, so it is plausible that it is really the set $\{x_\alpha\}$ described in Section 1. What we will do is show that M is a chain, which is again plausible if M is in fact $\{x_\alpha\}$. By (ii), M will have a largest element m. And by (i), this largest element will satisfy $m^+ = m$, so that it will be the desired maximal element of X.

4. The proof

As we noted in Section 2, it is enough to prove the special case of Zorn's lemma stated in Theorem 2. Thus we start with an arbitrary poset X, and we try to prove that the poset $\mathcal{X}$ of chains in X has a maximal element. We will apply the strategy described in Section 3 to $\mathcal{X}$.

We need a successor operation. If C is a nonmaximal element of $\mathcal{X}$, then there is a chain in X bigger than C, so we can choose $x \in X \smallsetminus C$ such that $C \cup \{x\}$ is a chain; set $C^+ := C \cup \{x\}$. [Note: We have used the axiom of choice.] If C is maximal, set $C^+ := C$. We also need a canonical upper bound of any chain in $\mathcal{X}$. But this is easy, since the union of a chain in $\mathcal{X}$ is again a chain, as we saw in the proof that Theorem 1 implies Theorem 2. Consider subsets $\mathcal{N} \subseteq \mathcal{X}$ with the following closure properties:

(i) If $C \in \mathcal{N}$, then $C^+ \in \mathcal{N}$.
(ii) If $\mathcal{C}$ is a chain in $\mathcal{N}$, then $\bigcup_{C \in \mathcal{C}} C$ is in $\mathcal{N}$.

Call $\mathcal{N}$ *closed* if it satisfies (i) and (ii). Note that $\mathcal{X}$ itself is closed, and the intersection of any family of closed sets is closed. In particular, the intersection of *all* closed subsets of $\mathcal{X}$ is closed. Call it $\mathcal{M}$; it is then the smallest closed set.

As explained in the previous section, the theorem will follow if we can show that $\mathcal{M}$ is a chain. Call an element $C \in \mathcal{M}$ *comparable* if it is comparable to every

$D \in \mathcal{M}$, i.e., $D \subseteq C$ or $C \subseteq D$. To prove $\mathcal{M}$ is a chain, we must show that every element of $\mathcal{M}$ is comparable. The following two lemmas are plausible in view of the intuition about what $\mathcal{M}$ really is.

Lemma 1. *Suppose C is comparable. If $D \in \mathcal{M}$ and $D \subsetneqq C$, then $D^+ \subseteq C$.*

Proof. Suppose not. Then $C \subsetneqq D^+$. But then $D \subsetneqq C \subsetneqq D^+$, contradicting the fact that D^+ was constructed by adjoining a single element of X to D. □

Lemma 2. *Suppose C is comparable. For every $D \in \mathcal{M}$, either $D \subseteq C$ or $D \supseteq C^+$.*

Proof. let $\mathcal{N} := \{D \in \mathcal{M} \mid D \subseteq C \text{ or } D \supseteq C^+\}$. We wish to show that $\mathcal{N} = \mathcal{M}$. Since $\mathcal{M}$ is the smallest closed set, it suffices to show that $\mathcal{N}$ is closed. Given $D \in \mathcal{N}$, we have $D \subsetneqq C$, $D = C$, or $D \supseteq C^+$. In the first case, $D^+ \subseteq C$ by Lemma 1, so $D^+ \in \mathcal{N}$; in the other two cases, $D^+ \supseteq C^+$, so again $D^+ \in \mathcal{N}$. This proves the first closure property. Next, suppose $\mathcal{C}$ is a chain in $\mathcal{N}$, and let $E := \bigcup_{D \in \mathcal{C}} D$. If every $D \in \mathcal{C}$ is a subset of C, then $E \subseteq C$, so $E \in \mathcal{N}$. Otherwise, some $D \in \mathcal{C}$ contains C^+; then $E \supseteq C^+$, and again $E \in \mathcal{N}$. This proves the second closure property. □

Now consider the set of comparable elements of $\mathcal{M}$. We will prove that this set is closed, hence is all of $\mathcal{M}$; this will complete the proof of the Hausdorff maximal principle. If C is comparable and $D \in \mathcal{M}$, then we know from Lemma 2 that either $D \subseteq C$ or $C^+ \subseteq D$. In either case D is comparable to C^+, so C^+ is a comparable set. Next, suppose $\mathcal{C}$ is a chain of comparable sets, and let $D := \bigcup_{C \in \mathcal{C}} C$. Given $E \in \mathcal{M}$, either $C \subseteq E$ for all $C \in \mathcal{C}$, in which case $D \subseteq E$, or else $E \subseteq C$ for some $C \in \mathcal{C}$, in which case $E \subseteq D$. Thus D is comparable, so the set of comparable sets is indeed closed.

The proof of the theorem is complete. As I've explicitly pointed out, the proof made use of the axiom of choice. This was used to define the successor operation. I mention this because the axiom of choice has been controversial historically. Nowadays, I think most mathematicians accept it but prefer to minimize its use because of its nonconstructive nature. (It is the unique axiom of set theory that asserts the existence of a set without describing the set explicitly.)

Exercises

1. If G is a finitely generated group and H is a proper subgroup, prove that H is contained in a maximal proper subgroup. Give an example to show that "finitely generated" cannot be deleted.

2. Show that one cannot eliminate the use of the axiom of choice in the proof of Zorn's lemma, because Zorn's lemma in fact implies the axiom of choice. [Hint: Consider partially defined choice functions suitably ordered, and use Zorn's lemma to prove the existence of a maximal one. Then show that this maximal one is in fact globally defined.]