Mathematics 4340
Zorn's lemma
Ken Brown, Cornell University, March 2009

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.*

The following special case is called the *Hausdorff maximal principle*:

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

To see why this is a special case, let X be an arbitrary poset and let 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 this special case. 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. The proof I will give in this handout is adapted from the proofs in Lang, *Real and Functional Analysis*, and Halmos, *Naïve Set Theory*.

1. Intuition behind the proof

For the purpose of explaining intuitively why Zorn's lemma is true, it will be convenient to make two simplifying assumptions. First, we assume that every chain in X actually has a *least* upper bound in X. (This is actually the case in most naturally occurring examples.) Second, we assume that we are given 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 intuitive proof now goes as follows:

Start with the smallest element $x_0 \in X$ (least upper bound of the empty chain). If it is maximal, we're done. Otherwise, set $x_1 := x^+ > x_0$. If x_1 is maximal, we're done. Otherwise, set $x_2 := x_1^+ > x_1$. Continuing *ad infinitum*, we either reach a maximal element or we generate an infinite chain

$$x_0 < x_1 < x_2 < \cdots .$$

In the latter case, let x_ω be the least upper bound of this chain. If it's maximal, we're done. Otherwise proceed as above to extend the chain to

$$x_\omega < x_{\omega+1} < x_{\omega+2} < \cdots .$$

Again, the process either terminates with a maximal element or leads to a new upper bound $x_{\omega+\omega} = x_{2\omega}$. We continue indefinitely until we reach a maximal element. This might take a long time. For instance, we might have to go to

$$x_\omega < \cdots < x_{2\omega} < \cdots < x_{3\omega} < \cdots < x_{\omega^2}$$

or even beyond.

Readers familiar with ordinal numbers and transfinite induction will see how to make this intuitive argument rigorous. [We are defining by transfinite induction an embedding $\alpha \mapsto x_\alpha$ of an initial segment of the ordinal numbers into X. The process has to terminate because there are too many ordinal numbers to fit them all into any set X.] To avoid a long excursion into the theory of ordinal numbers, however, I will use a surreptitious method of constructing the chain $\{x_\alpha\}$.

2. Strategy of the proof

Instead of building up the chain $\{x_\alpha\}$ 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 least 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 informally 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.

3. The proof

As we noted at the beginning, 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 2 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 $C^+ := C \cup \{x\}$. If C is maximal, set $C^+ := C$. 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 $D \subsetneq C$, then $D^+ \subseteq C$.*

Proof. Suppose not. Then $C \subsetneq D^+$. But then $D \subsetneq C \subsetneq D^+$, 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. Astute readers will have noticed that the proof made use of the axiom of choice. [See p. 909 of your text for a statement of that axiom.] This occurred at the very beginning of the proof, when the successor operation was defined. 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.

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.]