Math 146 Tutorial

March 2 – Zorn's Lemma

Zorn's Lemma is a foundational, set-theoretic tool used all the time in math. Although it's a somewhat strange and maybe counterintuitive result, it turns out to be *equivalent* to the Axiom of Choice: you can prove Zorn's Lemma using the Axiom of Choice, and vice-versa.

Partially ordered sets

Let X be some set and let $\leq$ be a relation; that is, for two elements $a, b \in X$, the statement $a \leq b$ is either true or false. We say that $\leq$ is a **partial order** on X if it has the following three properties:

(i) **Reflexivity:** $a \leq a$ for all $a \in A$.

(ii) **Antisymmetry:** If $a \leq b$ and $b \leq a$, then $a = b$.

(iii) **Transitivity:** If $a \leq b$ and $b \leq c$ then $a \leq c$.

The pair $(X, \leq)$ is called a **partially ordered set**. Notation and terminology are frequently simplified by just saying that X is a partially ordered set.

If you're worried about how to define this more carefully, you can think of $\leq$ as a subset of $X \times X$, and $a \leq b$ simply denotes that (a, b) is in this subset.

Example 1. $\mathbf{Z}$ has its usual ordering, *i.e.* $a \leq b$ if and only if $b - a \in \mathbf{N}$. This makes sense assuming we know how to define $\mathbf{N}$. One checks — and did check in Math 145! — that this satisfies properties (i), (ii), (iii) above, and is thus a partial ordering on $\mathbf{Z}$.

Example 2. Let $\mathbf{N}$ be the natural numbers ($0 \notin \mathbf{N}$), and define $a \leq b$ to mean $a|b$ for natural numbers $a, b \in \mathbf{N}$. This is a partial order on $\mathbf{N}$. We could have also defined $a \leq b$ to mean $b|a$, that would've worked too.

Example 3. Let $\mathbf{Z}[x]$ denote the set of polynomials with integer coefficients; for two such polynomials f, g, define $f \leq g$ to mean $\deg f \leq \deg g$. This satisfies (i) and (iii). But it doesn't satisfy (ii): for example, if $f(x) = x^2$ and $g(x) = -x^2$, then $\deg f = 2 = \deg g$ but $f \neq g$. The problem is that $\deg : \mathbf{Z}[x] \to \mathbf{N}_{\geq 0}$ is not injective.

Example 4. Let $\mathcal{P}$ be the collection of subsets of $\{1, 2, 3\}$, called the **power set** of $\{1, 2, 3\}$. Thus

$$\mathcal{P} = \Big\{\varnothing, \{1\}, \{2\}, \{3\}, \{1,2\}, \{1,3\}, \{2,3\}, \{1,2,3\}\Big\}.$$

Inclusion defines a partial order on $\mathcal{P}$: define $A \leq B$ to mean that $A \subseteq B$, *i.e.* $a \in A$ implies $a \in B$.

Let $A := \{2, 3\}$ and $B := \{1, 3\}$. Note that $A \not\leq B$ and $B \not\leq A$! Thus in a partial ordered set, it's *very possible* that two given elements are "incomparable" — neither $a \leq b$ nor $b \leq a$ are true.

A partially ordered set X is called **totally ordered** if, given $a, b \in X$, either $a \leq b$ or $b \leq a$. (Note that if both are true, then $a = b$ by (ii).)

Example 5. Thus $\mathbf{Z}$ is totally ordered and $\mathcal{P}$ is not. Despite this, $\mathcal{P}$ has many totally ordered *subsets*: for example

$$\Big\{\varnothing, \{2\}, \{2,3\}, \{1,2,3\}\Big\}$$

is totally ordered, since any two of its elements can be compared.

Let X be a partially ordered set and let Y be a subset of X. We say that Y is a **chain** in X if, for any $a, b \in Y$, either $a \leq b$ or $b \leq a$. So a chain is just a totally ordered subset.

Example 6. Let $\mathcal{C}[0,1]$ denote the set of continuous functions $f : [0,1] \to \mathbf{R}$. Define $f \leq g$ to mean $f(t) \leq g(t)$ for all $t \in [0,1]$. This is a partial ordering but not a total ordering, for example compare $f = \sin$ and $g = \cos$. We do have a chain though: let $f_n \in \mathcal{C}[0,1]$ be the function $f_n(t) := t^n$. Then

$$\{f_1, f_2, f_3, \ldots\}$$

is a chain in $\mathcal{C}[0,1]$, since $f_1 \geq f_2 \geq f_3 \geq \cdots$. (Check!)

Statement and discussion of Zorn's Lemma

Some intuitive terminology first. Let X be a partially ordered set and let Y be a subset of X. An **upper bound** for Y is an element $x_0 \in X$ such that $x_0 \geq y$ for all $y \in Y$. We say that Y is **bounded** if it has an upper bound.

Fix some $x_0 \in X$. We say that x_0 is **maximal** in X if $x_0 \not\leq x$ for all $x \in X \setminus \{x_0\}$. In other words, the only way $x_0 \leq x$ is if $x = x_0$. Think about this! It doesn't mean x_0 is bigger than everything; that would be an upper bound. Maximality means x_0 is *not smaller* than anything.

Example 7 (upper bounds)**.** A subset may or may not contain its upper bound. For instance consider $X = \mathbf{R}$ with the usual ordering and try $Y = (0,1)$. Then Y has upper bound 1, but $1 \notin Y$; in fact 1 is the **least** upper bound.

Or try $X = \mathcal{P}(\{1,2,3\})$ and $Y = \{\{1\},\{2\}\}$. Upper bounds include $\{1,2,3\}$ or $\{1,2\}$, but neither of these are in Y.

Example 8 (maximal elements)**.** Sometimes an upper bound is a maximal element. For example in $\mathcal{P}(\{1,2,3\})$, the element $x_0 := \{1,2,3\}$ is maxi*mum*, hence maximal. Maximum elements are always maximal.

The converse is not true. For example consider $X := \{\{1\},\{2\},\{3\}\}$, which is a subset of $\mathcal{P}(\{1,2,3\})$. Then $\{1\} \in X$ is maximal since nothing is bigger than it. But it's not maximum since it's not bigger than anything. [This X is pretty weird.]

Example 9. Some posets have no maximal element. Consider $\mathcal{C}[0,1]$ with ordering defined pointwise. If $f \in \mathcal{C}[0,1]$ then f is bounded, say $|f(x)| \leq M$ for all M. Then the constant function $g(x) := M + 1$ is strictly bigger than f at every point, so $f < g$. Therefore f is not maximal.

Now let $f_n(x) := nx$ for $n \geq 1$, so

$$X := \{f_1, f_2, f_3, \ldots\}$$

is a chain in $\mathcal{C}[0,1]$. It has no upper bound. Why? If $f \geq f_n$ for all n, then in particular $f(1) \geq f_n(1) = n$ for all n. This is impossible since $f(1) \in \mathbf{R}$. Therefore X is an unbounded chain.

Example 10 gives a poset in which has no maximal elements, but it also has an unbounded chain. Zorn's Lemma says that this is the only way a poset will fail to contain maximal elements.

Theorem 10 (Zorn's Lemma)**.** *Let X be a nonempty partially ordered set. If every chain in X is bounded, then X has a maximal element.*

The hypothesis of Zorn's Lemma is unnecessary.

Example 11. For instance let $X := \mathcal{P}(\mathbf{N}) \setminus \mathbf{N}$, so X is the collection of *proper* subsets of $\mathbf{N}$. We have an unbounded chain

$$\Big\{\{1\}, \{1,2\}, \{1,2,3\}, \dots\Big\}.$$

so the hypothesis of Zorn's Lemma is not active. But X lots of maximal elements, namely $\mathbf{N} \setminus \{k\}$ for any $k \in \mathbf{N}$. (In fact these are all the maximal elements.)

Application(s) of Zorn's Lemma

As I said earlier, Zorn's Lemma is used to prove all sorts of things in various branches. The strange aspect is that it is *not constructive* — it states that maximal elements exist under certain conditions, but does not tell you how to find such a maximal element. Despite this it has important theoretical uses. To namedrop a few:

- **Hahn–Banach Theorem:** If X is a complex normed vector space with subspace Y, then every bounded functional $\varphi : Y \to \mathbf{C}$ extends to a bounded functional on X.
- **Tychonoff's Theorem:** Let $\{X_\lambda\}$ be a family of compact topological spaces. Then the product topology on $\prod X_\lambda$ is compact.
- **Algebraic closures:** Let F be any field. Then F admits an algebraic closure, *i.e.* a field K containing F which is algebraically closed and algebraic over F.
- **Maximal ideals:** Let R be a commutative unital ring. Then every proper ideal in R is contained in a maximal ideal.
- **Bases in vector spaces:** Every vector space V over a field F has a basis.

Each of these are very important results, and their proofs make crucial use of Zorn's Lemma. Unfortunately most of them have inaccessible proofs for the moment, and their statements might mean nothing to you right now! In due time, my student ...

The last two are possible to handle right now though, and the third one is not too bad either. Let's prove the second last one.

Let R be a commutative ring with 1. You know from class that R is an R-module. Recall that an **ideal** in R is nothing but an R-submodule; in other words, an ideal is a subset I such that

(i) $0 \in I$,

(ii) $a, b \in I$ implies $a + b \in I$, and

(iii) $a \in I$, $r \in R$ implies $ra \in I$.

The set of all *proper* ideals forms a poset under inclusion. A **maximal idela** of R is just a maximal element of this poset.

Theorem 12. *Let R be a commutative ring with* 1 *and let I be a proper ideal of R. Then I is contained in a maximal ideal.*

Proof. We apply Zorn's Lemma. Let X be the set of proper ideals of R containing I; that is, it consists of ideals J such that $I \subseteq J \subsetneq R$. We have to show every chain has an upper bound.

Let Y be a (nonempty[1]) chain in X, and define

$$K := \bigcup_{J \in Y} J.$$

We claim that $K \in X$ and K is an upper bound for the chain Y. Let us verify this:

- We have to show that K is an ideal.
  - (i) Indeed 0 is in any $J \in Y$ and so $0 \in K$.
  - (ii) If $a, b \in K$ then $a \in J$ and $b \in J'$ for some $J, J' \in Y$; since Y is a chain we have $J \subseteq J'$ or $J' \subseteq J$. Wlog it's the first case. But then $a, b \in J'$, so $a + b \in J'$ since J' is an ideal, and so $a + b \in J' \subseteq K$.
  - (iii) Finally let $a \in K$, $r \in R$. Then $a \in J$ for some $J \in Y$, so $ra \in J \subseteq K$ as required.
- To show $K \in X$ we have to show K is *proper* and K contains I. But every $J \in Y$ contains I, so K contains I too.

  It's harder to show K is proper. Suppose not: $K = R$. Then in particular $1 \in K$, so $1 \in J$ for some $J \in Y$. But this is bad. Why? It implies that, for any $r \in R$, we have $r = r1 \in J$, and therefore $J = R$. So J is not proper, contradiction!
- Clearly K contains every $J \in Y$, so K is an upper bound for Y.

We've thus shown that every chain in X has an upper bound. By Zorn's Lemma, X admits an maximal element, say M. We'll show that M is a maximal ideal. Suppose not: $M \subsetneq M'$ for some proper ideal M'. But then $M' \in X$ clearly, contradicting the maximality of M as an element of X. ∎

The Axiom of Choice

Strictly speaking, I shouldn't have labeled Zorn's Lemma as a "Theorem" above — you can only consider it a theorem if you have an axiom by which you can prove it. It's true that if you Zorn's Lemma can be proven assuming the Axiom of Choice, but this is quite difficult. On the other hand, the Axiom of Choice can actually be proven using Zorn's Lemma!

Theorem 13 (Axiom of Choice)**.** *Let $\mathcal{X}$ be a collection of nonempty sets* — i.e. *its elements are nonempty sets. Then there exists a function $f : \mathcal{X} \to \bigcup_{S \in \mathcal{X}} S$ such that $f(S) \in S$ for all $S \in \mathcal{X}$.*

[1] The empty chain in a nonempty poset always has an upper bound.

We think of f as a "choice function". Why? The condition $f(S) \in S$ says that f takes a set S and gives you some element $x_s := f(S)$ belonging to S. So f is effectively *choosing* one element out each $S \in \mathcal{X}$.

Example 14. If $\mathcal{X}$ is the collection of nonempty subsets of $\mathbf{N}$, we can define f easily — let $f(S)$ be the least element of S. This makes sense since $\mathbf{N}$ is well-ordered.

Example 15. Let $\mathcal{X}$ be the collection of open intervals (a, b) in the real line, with $a < b$. This is a huge uncountable set. For each such interval (a, b)

$$f((a, b)) := \frac{a+b}{2} \in (a, b).$$

Thus f is a choice function.

Example 16. Let $\mathcal{X}$ be the collection of *all* subsets of $\mathbf{R}$, not necessarily intervals as in Example 15. How can we define a choice function? We were able to write down a formula in Example 15 because we knew we were dealing with intervals. But that method won't work here.

So we've seen that in some cases we may define choice functions manually, but in other cases it can be quite elusive. Nonetheless, the Axiom of Choice states that choice functions *always* exist, and we can prove it with Zorn's Lemma.

Proof of Theorem 14 using Theorem 11. Let $\mathcal{X}$ be a collection of nonempty sets. For each subcollection $\mathcal{Y}$, let $\mathscr{Z}$ consist of all choice functions $f : \mathcal{Y} \to \bigcup_{S \in \mathcal{Y}} S$ such that $f(S) \in S$ for all $S \in \mathcal{Y}$, and let

$$\mathscr{Z} := \bigcup_{\mathcal{Y} \subseteq \mathcal{X}} \mathscr{Z}_{\mathcal{Y}}.$$

So $\mathscr{Z}$ — Z for "Zorn" — consists of all choice functions defined on *subcollections* of $\mathcal{X}$. Our goal is to show $\mathscr{Z}_{\mathcal{X}}$ to be nonempty. We'll do this by applying Zorn's Lemma to find a maximal element of $\mathscr{Z}$, and show that this must be an element of $\mathscr{Z}_X$. Notation: for $f \in \mathscr{Z}_{\mathcal{Y}}$ let us denote $\operatorname{dom} f := \mathcal{Y}$.

Why is $\mathscr{Z}$ nonempty? Notice that the empty set has a choice function: namely, the "empty" function that takes the empty set and produces nothing. Strangely but conveniently, this counts as a function if you go back to the set-theoretic definition of function. So it's an element of $\mathscr{Z}_{\varnothing}$ and hence of $\mathscr{Z}$.

We make $\mathscr{Z}$ into a poset as follows: if $f, g \in \mathscr{Z}$,

$$f \leq g \iff \operatorname{dom} f \subseteq \operatorname{dom} g \text{ and } g(S) = f(S) \text{ for all } S \in \operatorname{dom} f.$$

To use Zorn's Lemma we have to show every chain has an upper bound. Indeed suppose we have a chain $\mathscr{C}$ — C for chain — and let $\mathcal{Y} := \bigcup_{f \in \mathscr{C}} \operatorname{dom} f$. Now define $g \in \mathscr{Z}_{\mathcal{Y}}$ as follows: for $S \in \mathcal{Y}$ we have $S \in \operatorname{dom} f$ for some $f \in \mathscr{C}$, thus set

$$g(S) := f(S).$$

Prove that this is well-defined — *i.e.* independent of f —, that $g \in \mathscr{Z}_{\mathcal{Y}}$, and that $g \geq f$ for all $f \in \mathscr{C}$. So g is an upper bound for the chain $\mathscr{C}$.

Therefore we can apply Zorn's Lemma: $\mathscr{Z}$ has a maximal element, say f. So f is a choice function on subsets of $\mathcal{Y} := \operatorname{dom} f$. We'll show that in fact $\mathcal{Y} = \mathcal{X}$, so f is choice function we're looking for. If $\mathcal{Y} \subsetneq \mathcal{X}$ then we can choose some set $S_0 \in \mathcal{X}$ which is not in $\mathcal{Y}$. But S_0 is nonempty, so we can find some $x_0 \in S_0$; define a new choice function $g \in \mathscr{X}_{\mathcal{Y} \cup \{S_0\}}$ by

$$g(S) := \begin{cases} f(S) & \text{if } S \in \mathcal{Y}, \\ x_0 & \text{if } S = S_0. \end{cases}.$$

Then indeed g is a choice function, and $f < g$. This contradicts maximality of f. ∎

Conversely, the Axiom of Choice (Theorem 14) can be used to prove Zorn's Lemma (Theorem 11).

Proof of Theorem 11 using Theorem 14. ... ∎