ONE WAY TO THINK ABOUT: ZORN'S LEMMA

Theorem 1 (Zorn's Lemma)**.** *Let $\mathcal{P}$ be a partially ordered set. If every chain in $\mathcal{P}$ has an upper bound, then $\mathcal{P}$ has a maximal element.*

To get comfortable with the statement, let's ensure we are familiar with the notions of *partially ordered sets*, *chains*, *upper bounds*, and *maximal elements*.

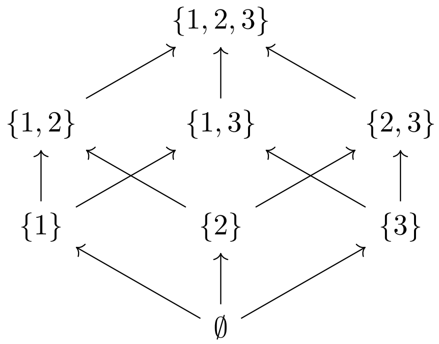
1. Definitions

Definition 2. *A set $\mathcal{P}$ is partially ordered by $\preceq$ if the following conditions hold:*

(1) *if $x \preceq y$ and $y \preceq z$ then $x \preceq z$;*
(2) *$x \preceq x$;*
(3) *if $x \preceq y$ and $y \preceq x$ then $x = y$*

for all $x, y, z \in \mathcal{P}$.[1]

For instance, a common example is the ordering $\subseteq$ on the finite set $\{1, 2, 3\}$, as the following *Hasse diagram* shows; each arrow symbolises set inclusion. Formally, here $\mathcal{P} = \mathcal{P}(\{1, 2, 3\})$ (the power set of $\{1, 2, 3\}$), and $\preceq$ is $\subseteq$.



Note that not all elements of the power set of $\{1, 2, 3\}$ are comparable. Indeed, $\{1\} \not\subseteq \{2\}$ and $\{2\} \not\subseteq \{1\}$. Partial orders where all pairs of elements are comparable are called *linear orders*.

Definition 3. *A subset $C \subseteq \mathcal{P}$ of a partial order is called a* chain *if it is linearly ordered under the partial ordering of $\mathcal{P}$.*

For instance, $\{\emptyset, \{1\}, \{1, 2, 3\}\}$ is a chain in $\mathcal{P}(\{1, 2, 3\})$, while $\{\{2\}, \{1, 3\}\}$ is not.

Let $(\mathcal{P}, \preceq)$ be a partial ordering. The following two definitions are very important.

Definition 4. *Let $C \subset \mathcal{P}$ be a chain. An element $b \in \mathcal{P}$ is called an* upper bound of C *if*

$$\textit{whenever } x \in C \textit{ then } x \preceq b.$$

So every element of C is below b under the ordering $\preceq$ of $\mathcal{P}$.

Definition 5. *An element $m \in \mathcal{P}$ is called* maximal *if*

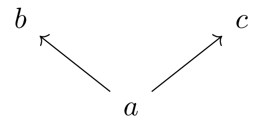
$$\textit{whenever } x \in \mathcal{P} \textit{ then } m \not\preceq x.$$

So no element of $\mathcal{P}$ is above m under the ordering $\preceq$ of $\mathcal{P}$.

Remark 1. *Observe that maximal elements need not be unique, and that they need not be above every element. For instance, in the ordering given by the Hasse diagram*

May 31, 2021 (linus.richter@sms.vuw.ac.nz).

[1] These properties are called *transitivity*, *reflexivity*, and *anti-symmetry*.



the element b is maximal (no element is above it), even though c is not below b; and for the same reason, c is maximal, too.

Now we can make sense of the statement of Zorn's Lemma: all we need to do to show that a maximal element exists is:

(1) define the *correct* partial ordering (in the sense that it provides the maximal element we are looking for); and
(2) show that every chain has an upper bound.

Remark 2. *Ensure that you know the difference between upper bounds and maximal elements; and that they are not necessarily (but could be) equal!*

Here are some examples, following this template.

2. Applications

The key to Zorn's Lemma is twofold. On the one hand, we need to define the chains properly, in order for what it means for an element to be maximal to agree with the notion of maximality we require. Secondly, we need to actually verify that each chain has an upper bound.

Application 1. *Every non-empty finitely generated non-trivial proper ideal of a commutative ring contains a proper maximal subideal.*

Proof. Let R be a commutative ring. Suppose $I = (a_0, \ldots, a_n)$ is a proper subideal of R. It is also non-trivial: $I \neq \{0\}$. We want to find a proper ideal $J \subset R$ that is maximal in I. Here, by maximal we mean that

$$(*) \qquad \text{whenever } J \subset K \subset I \text{ and } K \text{ is an ideal in } R \text{ then } K = J \text{ or } K = I.$$

We follow the template.

<u>Step 1</u>: Recall that we are looking for a proper maximal subideal, where maximal ideals are defined as above. Consider the partial ordering

$$\mathcal{P} = \{K \subset R \mid K \text{ is an ideal in } R \text{ and } K \text{ is a proper subset of } I\}$$

under set inclusion. Before we continue, let's ensure that a maximal element in $\mathcal{P}$ is exactly the type of maximal element we need according to $(*)$.

Claim 1. An ideal J satisfies $(*)$ if and only if it is maximal with respect to $\mathcal{P}$.

Proof of Claim 1. If J satisfies $(*)$ then there is no ideal K' in R such that $K \subset K' \subset I$ and all inclusions are proper. Hence there is no element above J in the ordering of $\mathcal{P}$ (since each such element would witness failure of $(*)$).

Now suppose J is maximal in $\mathcal{P}$. So there is no element of $\mathcal{P}$ above J, or in other words, there is no ideal properly between J and I. Hence $(*)$ holds. $\dashv$

<u>Step 2</u>: We now verify that each chain has an upper bound. Suppose $C \subset \mathcal{P}$ is a chain (so any two elements in C are comparable). If C is empty, just pick any element of $\mathcal{P}$ and we are done. This is possible since I is non-trivial, so $I \neq \{0\}$. Yet $\{0\} \in \mathcal{P}$.

Now assume C is non-empty. Consider the union of elements of C, denoted by L. Hence each element of C, i.e. each ideal in C, is a subset of L. Then L is also an ideal of R.[2]

Claim 2. Every element of C is below L. In other words, L is an upper bound of C.

Proof of Claim 2. Follows from the definition of the ordering of $\mathcal{P}$. $\dashv$

[2] Convince yourself that this is true.

That's a good start, but we also need to make sure that $L \in \mathcal{P}$. Since L is an ideal in R, it suffices to ensure that $L \neq I$.

Claim 3. $L \neq I$.

Proof of Claim 3. Suppose $L = I$. Hence $L = (a_0, \ldots, a_n)$, and so each $a_i \in L$. Recall that L is the union of ideals in C, so each a_i is contributed to L by some K_i. Since C is a chain, all of its elements are linearly ordered under the ordering of $\mathcal{P}$, i.e. under set inclusion. And since I is finitely generated, there are only finitely many ideals K_i needed to contribute all a_i to L. For example, assume without loss of genrality

$$K_0 \subseteq K_1 \subseteq \ldots \subseteq K_n.$$

Then $\{a_0, \ldots, a_n\} \subseteq K_n$, and since K_n is an ideal, we have $(a_0, \ldots, a_n) \subseteq K_n$.But now $I \subseteq K_n$, contradicting that $K_n \in \mathcal{P}$. Hence $L \neq I$. ⊣

So we have verified all the conditions of Zorn's Lemma, and we may now apply it. By Zorn's Lemma, there is a maximal element $J \in \mathcal{P}$. And by Claim 1, this J is exactly as needed. □

Observe how Zorn's Lemma does not give us a maximal subideal directly; it just tells us that one exists. Zorn's Lemma does also not assert that the maximal element provided is unique; there can be many maximal elements!

Another way of thinking about Zorn's Lemma is this: we try to construct a very large object by showing that we can merge (arbitrarily many) coherent smaller building blocks. The following example, from graph theory, uses this point of view.

Application 2. *Every connected graph has a spanning tree.*

Proof. Recall some notions from graph theory: a *walk* is a sequence of edges. A *cycle* is a walk that starts and ends at the same vertex. A *path* is a walk where all edges and vertices are distinct. A graph is *connected* if for any two vertices there exists a path connecting them. A subgraph T of a graph G is called a *tree* if it is connected and has no cycles. A *spanning tree* is a tree that contains all vertices.

We want to use the template. What we need is (1) the correct partially ordered set, and (2) a proof that Zorn's Lemma applies.

Step 1: We try to build a "large" tree, so let's consider its building blocks: define

$$\mathcal{P} = \{T \subseteq G \mid T \text{ is a tree in } G\}$$

ordered by inclusion. Now, a maximal element m in $\mathcal{P}$ is a tree such that any subgraph of G that properly contains m is not a tree. Let's verify that this notion of maximal agrees with our required notion:

Claim 1. A maximal element of $\mathcal{P}$ is a spanning tree of G.

Proof of Claim 1. Let m be maximal in $\mathcal{P}$ (so, in particular, m is a tree). Let $v \in G$ be a vertex. For contradiction, suppose v does not lie on m. Let p be a path ending at v that meets m. Such a path exists since G is connected. Since every path is finite, so is p; hence there is a least vertex v' on m such that p', the path p starting at v' (i.e. omitting all edges and vertices of p up to v') does not contain any vertices on m bar v'. But then m extended by p' is a tree properly extending m, and hence m is not maximal, a contradiction. ⊣

Step 2: So now it suffices to show that every chain in $\mathcal{P}$ has an upper bound; but that is easy to show: take the union of elements of the chain. That will again be a tree (if it was cyclic, then a cycle would have already existed in an element of the chain, which contradicts the fact that the chain is a subset of $\mathcal{P}$). Hence Zorn's Lemma applies, and the claim proves there is a spanning tree. □

Here is another example, in less detail.[3]

Application 3. *Every vector space has a basis.*

Proof. Let V be a non-empty vector space. Consider the partially ordered set

$$\mathcal{P} = \{B \subset V \mid B \text{ is linearly independent}\}$$

ordered under set inclusion.

A maximal element in $\mathcal{P}$ is a linearly independent subset of V that cannot be extended. But even more is true:

Claim 1. A maximal element in $\mathcal{P}$ is a basis for V.

Proof of Claim 1. Let m be maximal in $\mathcal{P}$. If m is not a basis for V then there is $v \in V$ such that $v \notin \operatorname{span}(m)$. But then $m \cup \{v\}$ is linearly independent. Hence $m \cup \{v\} \in \mathcal{P}$, and $m \subset m \cup \{v\}$. Thus m is not maximal, a contradiction. ⊣

Now there is only one thing left to check:

Claim 2. Every chain in $\mathcal{P}$ has an upper bound.

Proof of Claim 2. Let C be a chain in $\mathcal{P}$. Then consider the union of C, denoted by L. Suppose L is not linearly independent. Then there are vectors $v_1, \ldots, v_n$ such that

$$\lambda_1 v_1 + \ldots + \lambda_n v_n = 0$$

and not all λ_i are zero. Just like in Application 1, the finitely many vectors v_i are contributed by finitely many $B_i \in C$. Hence, the largest of all of them is already linearly dependent, contradicting that each B_i is an element of $\mathcal{P}$ and hence linearly independent. So $L \in \mathcal{P}$, and the claim is true. ⊣

So Zorn's Lemma applies, and so $\mathcal{P}$ has a maximal element; this is a basis by Claim 1. □

Note that the argument above holds for all vector spaces, no matter whether finite- or infinite-dimensional. But also observe, that Zorn's Lemma doesn't actually construct a basis; it only asserts that one exists.[4]

3. Recap

We need to do two things when we want to use Zorn's Lemma:

(1) Define a partially ordered set whose maximal elements satisfy exactly the property we are looking for; and
(2) check whether Zorn's Lemma is actually applicable, by checking that every chain has an upper bound in the partial ordering.

If the maximal elements of the partially ordered set don't have the property we require, a different ordering might be needed.

[3] This result cannot be proven without Zorn's Lemma, or an equivalent result; check out the *Axiom of Choice* if you think this is interesting. In fact, the statement "every vector space has a basis", Zorn's Lemma, and the Axiom of Choice are logically equivalent!

[4] The *Axiom of Choice* gives many results of this type, and existence proofs where objects are constructed using AC are usually called non-constructive existence proofs.