# Curio 2: Infinite-dimensional vector spaces and Zorn's lemma

mike miller eismeier

## Preamble

When submitting solutions to this (or any other) Curio, please write solutions to the things labeled **Problem.** below. Here, there are three Problems, and the first has four parts. When actually working through it, you should read and attempt to understand all linked references and all discussion.

Please take your time and be careful in writing up your solutions, because the Curio submission is a not insubstantial portion of your final grade.

I advise reading each Curio and maybe trying to think through the problems, but only submitting formally written solutions halfway into the term.

This Curio is highly dependent on the work from Curio 1 (which you may take for granted while working on this, but should at least understand). Later curios will be more independent. The first parts of this are expository; your work starts on the last two pages the document.

## Curio 2

In the notes and in class, we've developed the theory of *finite-dimensional vector spaces.* For a number of foundational lemmas, I used things which ultimately relied on an inductive argument (any time I referred to the notion of "redundancy", I ordered our vectors into a finite list).

Most of these make sense for infinite-dimensional vector spaces, except for the very final thing we covered in Chapter 4.7.

**Definition 1.** *Write $Map_{fin}(S, \mathbb{F})$ for the set of functions $a : S \to \mathbb{F}$ so that $a(v_s)$ is nonzero for only finitely many $s$. If $S \subset V$ is a set of vectors in $V$, this guarantees the expression*

$$\sum_{v_s \in S} a(v_s) v_s \in V$$

*makes sense, because only finitely many terms are nonzero.*

*In this case, we say a* ***linear relation*** *between the elements of $S$ is a function $a \in \mathrm{Map}_{\mathrm{fin}}(S, \mathbb{F})$ for which*

$$\sum_{v_s \in S} a(v_s) v_s = \vec{0}.$$

*The trivial linear relation is the function $a(v_s) = 0$ for all $v_s \in S$. We say the set $S$ is* ***linearly independent*** *if the only linear relation between the elements of $S$ is the trivial linear relation.*

The notion of 'basis' then means exactly what it has before: a linearly independent spanning set. (The equivalent statement is now "$S$ is a basis if and only if, for all $v \in V$, there exists a unique $a \in \mathrm{Map}_{\mathrm{fin}}(S, \mathbb{F})$ so that $v = \sum_{v_s \in S} a(v_s) v_s$; that is, every element can be written as a finite linear combination of elements of $S$ in a unique way.)

To make any progress thinking about these I need a way to carry out the iterative arguments I had in mind before, but with infinite sets. One way to do this is in terms of an extension to infinite sets called "transfinite induction". This is not my preferred method; I think that most such arguments are more clearly phrased in terms of "Zorn's lemma".

## Partially ordered sets

Zorn's lemma is a statement about partially ordered sets, so we need to define those.

**Definition 2.** *Given a set $P$, a* ***partial order*** $\leqslant$ *on $P$ is a rule which, given a pair $(x, y) \in P \times P$, asserts whether $x \leqslant y$ or $x \nleqslant y$. We demand that this rule satisfies the following three properties:*

- ***Reflexivity.*** *We have $x \leqslant x$ for all $x \in S$.*
- ***Antisymmetry.*** *If $x \leqslant y$ and $y \leqslant x$, then $x = y$.*
- ***Transitivity.*** *If $x \leqslant y$ and $y \leqslant z$, then $x \leqslant z$.*

*A partially ordered set is a set $P$ equipped with a partial order $\leqslant$.*

Formally, $\leqslant$ can be understood as a subset $\Gamma_{\leqslant} \subset P \times P$, of pairs $(x, y)$ for which $x \leqslant y$. I will not think in terms of this formalism. One important thing to note is that not all elements are comparable: you can have $x, y \in P$ for which neither $x \leqslant y$ nor $y \leqslant x$ is true. Before using this notion, it's worth giving an example, to show how different it is from the idea you're used to.

*Example* 1. Let $X$ be any set, and let $\mathcal{P}(X)$ be the power set of $X$, with the order $S \leqslant T$ if $S \subset T$. Then $S \subset S$, so this is reflexive; if $S \subset T$ and $T \subset S$ then $T = S$, so this is antisymmetric; and if $S \subset T$ and $T \subset U$ then $S \subset U$, so it is a transitive relation.

However, for instance, neither $\{1\}$ nor $\{2, 3\}$ are contained in the other, so these are incomparable elements of $\mathcal{P}(\{1, 2, 3\})$.

**Definition 3.** *A* ***maximum*** *of a partially ordered set $(S, \leqslant)$ is an element $x \in S$ so that for all $y \in S$, we have $y \leqslant x$.*

*A* ***maximal element*** *of a partially ordered set $(S, \leqslant)$ is an element $x$ so that* ***for all*** *$y$* ***comparable to*** *$x$, we have $y \leqslant x$.*

The restriction in the second part means that we only look at those $y$ for which at least one of $y \leqslant x$ or $x \leqslant y$ is true. An equivalent phrasing is "$x$ is maximal if **there does not exist** a $y \neq x$ for which $x \leqslant y$": there are no larger elements, even if there are other incomparable elements.

The definitions of minimums and minimal elements are similar.

*Example* 2. Write $\mathcal{P}_{\neq}(\mathbb{N})$ for the set of **proper** subsets of $\mathbb{N}$, equipped with the relation $S \leqslant T$ if $S \subseteq T$ (the containment relation, where $S \subseteq T$ means $x \in S$ implies $x \in T$; it is possible here that $S = T$). Then $\mathcal{P}_{\neq}(\mathbb{N})$ has a minimum, $\varnothing$; the empty set is contained in every other set!

The set $\{0, 2, 3, 4, 5, \cdots\} = \{1\}^c \subset \mathbb{N}$ is a proper subset of $\mathbb{N}$. There is only one larger subset of $\mathbb{N}$, which is $\mathbb{N}$ itself — not a proper subset, so not an element of $\mathcal{P}_{\neq}(\mathbb{N})$. Therefore, $\{1\}^c$ is a maximal element of $\mathcal{P}_{\neq}(\mathbb{N})$. However, it is not a **maximum**, because it does not contain eg $\{1\}$ — there is an element incomparable to it.

*Example* 3. Not every partially ordered set has a maximal element. For instance, $(\mathbb{N}, \leqslant)$ with the usual ordering has no maximal element: given $n$, the element $n + 1$ is always larger.

*Example* 4. Write $\mathsf{Sub}_S(V)$ for the set of subspaces of a vector space $V$ which contain a given subset $S \subset V$, and equip it with the relation $W \leqslant U$ if $W \subset U$. Our first theorem about spans asserts that $\text{span}(S)$ is the minimum in this partially ordered set.

We need to give two more definitions for the statement of Zorn's lemma.

**Definition 4.** *Given a partially ordered set $(P, \leqslant)$, a* ***chain*** *is a subset $C \subset P$ for which all elements are comparable.*

*Example* 5. For instance,

$$C = \big\{\{1\}, \{1, 2\}, \{1, 2, 3\}, \{1, 2, 3, 4\}, \cdots\big\} \subset \mathcal{P}(\mathbb{N})$$

is a chain. Each pair of sets $\{1, \cdots, n\}$ and $\{1, \cdots, m\}$ has one contained in the other (which it is depends on whether $m \leqslant n$ or $n \leqslant m$). Chains don't have to be countable sequences of elements like this, in general, but that is how people often imagine them: "increasing chains" instead of "increasing sequences".

**Definition 5.** *Given a subset $S \subset P$, an* ***upper bound*** *for $S$ is an element $x \in P$ so that, for all $s \in S$, we have $s \leqslant x$. (In particular, $x$ is comparable to every element of $S$.)*

## Zorn's lemma

We can finally state Zorn's lemma.

**Lemma 1** (Zorn's lemma)**.** *Suppose $(P, \leqslant)$ is a partially ordered set with the following property: for every chain $C \subset P$, there exists an element $x_C \in P$ which is an upper bound for $C$.*

*Then there exists a maximal element $x \in P$.*

It turns out that this statement is completely equivalent to the axiom of choice. I will give the proofs here to show you a test case of how one might use each of Zorn's lemma and the axiom of choice.

**Lemma 2.** *The axiom of choice follows from Zorn's lemma.*

*Proof.* Suppose we have a set $I$ parameterizing a family of sets $X_i$. I claim there exists a function $f : I \to \cup_I X_i$ so that $f(i) \in X_i$ for all $i$, and thus that there exists an element of $\prod_{i \in I} X_i$. This is called a "choice function for $I$".

To prove this, consider the partially ordered set

$$P = \{(A, f) \mid A \subset I \text{ and } f : A \to \cup_{i \in I} X_i \text{ is a choice function for } A\}.$$

These might be called "partial choice functions", where I have a choice function on a subset $A \subset I$. The order relation has $(A, f) \leqslant (A', f')$ if $A \subset A'$ and $f'(a) = f(a)$ for all $a \in A$.

I want to apply Zorn's lemma to $P$. First, I will show that any chain in $P$ has an upper bound. Suppose we have a chain $(A_c, f_c) \in P$ labeled by $c \in C$. Consider the set $A = \cup_{c \in C} A_c$ with the function $f : A \to \cup_{i \in I} X_i$ defined by

$$f(a) = f_c(a) \text{ if } a \in A_c.$$

Because this is a chain, whenever $a \in A_c$ and $a \in A_{c'}$, one of $(A_c, f_c)$ or $(A_{c'}, f_{c'})$ is less than or equal to the other, meaning that (eg) $f_{c'}(x) = f_c(x)$ for all $x \in A_c$. Because our $a$ lies in both of these sets, we see that $f_c(a) = f_{c'}(a)$, so that the function defined above is unambiguous (did not depend on a choice of $c$ for which $a \in A_c$).

Thus $(A, f)$ is defined, and $A_c \subset A$ for all $c \in C$ and whenever $a \in A_c$, we have $f(a) = f_c(a)$. Thus $(A_c, f_c) \leqslant (A, f)$ for all $c \in C$, so that this is indeed an upper bound for the chain.

Now Zorn's lemma implies that there exists a maximal $(A, f)$. I claim that $A = I$. For if $(A, f) \in P$ and $A \neq I$, it is not maximal: I can choose an element $i \in I$ for which $i \notin A$, and choose some $x_i \in X_i$. Then I define $(A', f')$ for which $A' = A \cup \{i\}$ and

$$f'(j) = \begin{cases} f(j) & j \in A \\ x_i & j = i \end{cases}.$$

This has $A \subset A'$ and the restriction of $f'$ to $A$ is $f$, so that $(A, f) \leqslant (A', f')$ and $(A', f') \neq (A, f)$, so indeed $(A, f)$ is not maximal.

Thus the maximal $(A, f)$ has $A = I$, in which case $(I, f) \in P$ means precisely that $f$ is a choice function for $I$, as desired. □

(Do you see how this feels a little bit like a generalized induction? The inductive step was where we said "the maximal element has to be everything", because I could just add one more element. But there's also one more step, in showing that every chain has an upper bound.)

**Lemma 3.** *Zorn's lemma follows from the axiom of choice.*

This is actually substantially harder, and is not worth going into for my purposes. If you want to read the proof, I recommend this handout. Let me point out an application I like (it tells me that I can compare the sizes of any two sets).

**Corollary 4.** *Assuming the axiom of choice, for any two sets $X$ and $Y$, there exists either an injection $f : X \to Y$ or an injection $g : Y \to X$. Thus either $|X| \leqslant |Y|$ or $|Y| \leqslant |X|$.*

*Proof.* Consider the set $P = \{(A, f) \mid A \subset X \text{ and } f : A \to Y \text{ injective}\}$, with the order $(A, f) \leqslant (A', f')$ when $A \subset A'$ and $f'(a) = f(a)$ for all $a \in A$.

If $(A_c, f_c)$ (labeled by $c \in C$) is a chain, then the same argument as above verifies that $(A, f) = (\cup A_c, \cup f_c)$ is well-defined. Notice that $f$ is still injective: if $x, y \in A$, then $x \in A_c$ and $y \in A_{c'}$ for some $c, c'$; because $C$ is a chain, one of $A_c \subset A_{c'}$ or $A_{c'} \subset A_c$ (say the latter for convenience), so that $x, y \in A_c$; then because $f_c$ is injective we see that $f(x) = f(y) \implies f_c(x) = f_c(y) \implies x = y$. So $f$ is injective, and $(A, f)$ is an upper bound for this chain.

Zorn's lemma implies that there is a maximal element in the set $P$. I claim that if either $A = X$ or $f$ is surjective. To prove this, I'll show that if $A \subset X$ is proper and $f$ is not surjective, $(A, f)$ is not maximal. To see this, pick some $y \notin f(A)$, and pick some $x \notin A$; then define $(A', f')$ to be $A' = A \cup \{x\}$ and

$$f'(a') = \begin{cases} f(a') & a' \in A \\ y & a' = x \end{cases}.$$

This new function is injective, defined on a larger domain, and restricts to $f$ on the original domain. Thus $(A, f)$ is not maximal in $P$.

Thus for our maximal $(A, f)$ we either have $A = X$ — in which case $f : X \to Y$ is an injection, as desired — or $(A, f)$ has $f : A \to Y$ surjective (and $f$ injective because $(A, f) \in P$ requires that $f$ is injective). In the latter case, $f : A \to Y$ is a bijection, and therefore has an inverse function $f^{-1} : Y \to A$. Composing this with the inclusion $i : A \to X$, we obtain an injective map $g : Y \to X$, as desired. □

## Basis restriction and existence

Now you're mostly on your own. I'm going to tell you the restatements of some facts from linear algebra, adapted to the infinite-dimensional context. For some, I will tell you what set to apply Zorn's lemma on; for others, I will ask you to figure out the whole strategy.

**Lemma 5** (Basis restriction lemma)**.** *Let $V$ be a vector space. If $S \subset V$ spans $V$, then there is a subset $S' \subset S$ which is a basis for $V$.*

**1.** Prove the previous lemma by applying Zorn's lemma to the set $P = \{S' \subset S \mid S' \text{ is linearly independent}\}$, with the containment relation.

**Proposition 6.** *Let $V$ be an arbitrary vector space. There exists a basis for $V$.*

**2.** Prove the previous proposition. (All you really have to do here is provide me a spanning set for $V$.)

This statement is in fact equivalent to the axiom of choice (though I will not prove it). The bases you produce are completely inexplicit; you will never get a handle on them.

**Corollary 7.** *There exists a basis for $\mathbb{R}$ as a vector space over $\mathbb{Q}$. This means there exists a set $S \subset \mathbb{R}$ so that every real is a rational linear combination of elements of $S$ in a unique way. One may even choose $S \subset [0, 1]$.*

You will never be able to tell me such a set $S$; it can only be constructed in this highly inexplicit fashion. It is used in measure theory to construct a non-measurable subset of $\mathbb{R}$ (a "Vitali set").

## Size of linearly independent sets

I would like to say there is a good notion of "dimension" of infinite-dimensional vector spaces.

**Proposition 8.** *Suppose $V$ is a vector space. If $S \subset V$ spans $V$, and $I \subset V$ is linearly independent, then there exists an injection $f : I \to S$, and thus $|I| \leqslant |S|$.*

The strategy is analogous to the finite-dimensional case, but uses Zorn's lemma. Consider the set

$$P = \{(I', f) \mid I' \subset I \text{ and } f : I' \to S \text{ an injection, and } I' \cup (S \backslash f(I')) \text{ spans } V\}.$$

Here recall the notation $S \backslash f(I')$ to mean

$$S \backslash f(I') = \{s \in S \mid s \notin f(I')\};$$

it is the complement of $f(I')$ in the set $S$. The idea is that we are replacing elements of $S$ with elements of $I'$ without changing the size of the spanning set.

**3.** Give a careful proof of the previous Proposition.

**Corollary 9.** *Every vector space has a well-defined dimension* $\dim(V)$*, the cardinality of a basis for* $V$*. If* $W \subset V$*, we have* $\dim(W) \leqslant \dim(V)$*.*

*Proof.* If $S$ and $S'$ are two bases for $V$, the previous Proposition produces an injection $f : S \to S'$ and an injection $g : S' \to S$. The Cantor-Schroeder-Bernstein theorem asserts that when this is the case, there exists a bijection $h : S \to S'$, so that $|S| = |S'|$. We set $\dim(V) = |S|$ for some basis $S$ of $V$. □

In particular, $\dim_{\mathbb{Q}}(\mathbb{R}) = |\mathbb{R}|$ is an uncountable-dimensional vector space; the basis we constructed above has uncountably many elements.

To conclude, recall that for finite-dimensional vector spaces, proper subspaces have strictly smaller dimension. This is false for infinite-dimensional vector spaces, and as a result most of our results about finite-dimensional vector spaces will not extend to the general case.

**4.** Consider the vector space $\mathbb{F}^\infty$ of eventually-zero sequences in $\mathbb{F}$. Show that there exists a proper subspace $V \subset \mathbb{F}^\infty$ of the same dimension as $\mathbb{F}^\infty$, meaning that there is a basis $S$ for $\mathbb{F}^\infty$, a basis $S'$ for $V$, and a bijection $f : S \to S'$.