Appendix: Zorn's Lemma

At a couple points during the course we have benefited from Zorn's Lemma. Here we outline the statement of Zorn's Lemma and give an example of its use.

Let S be a set. A **partial order** on S is a binary relation $\leq$ such that for all $x, y, z \in S$,

(i) $x \leq x$;
(ii) if $x \leq y$ and $y \leq x$, then $x = y$; and
(iii) if $x \leq y$ and $y \leq z$, then $x \leq z$.

Note that there can be many pairs of elements $\{x, y\}$ in X which are incomparable, i.e. $x \not\leq y$ and $y \not\leq x$. A **chain** is a subset $C \subseteq x$ such that for all $x, y \in S$, either $x \leq y$ or $y \leq x$. We write $x < y$ as an abbreviation for the statement that '$x \leq y$ and $x \neq y$'. If $S \subseteq X$, an *upper bound* for S is an element $b \in X$ such that $s \leq b$ for all $s \in S$. We say that S is **bounded above** if such an upper bound for S exists. Note that b is *not* required to belong to the subset S in this case. A **maximal element** in X is an element $m \in X$ such that no element of X is larger than m; that is, there does not exist $x \in X$ such that $m < x$.

Example: $\mathbb{Z}$ with Divisibility. An example is the relation of divisibility on the set of integers, in which the pair $\{4, 15\}$ is incomparable since $4 \nmid 15$ and $15 \nmid 4$. In this setting, $\{1, 2, 4, 8, 16, \ldots\}$ is a chain with no upper bound. The chain $\{3, 12, 36, 1440\}$ has many choices of upper bound: 1440 is an upper bound (the *least* upper bound), and 2880 is also an upper bound. There is no maximal element in $\mathbb{Z}$ for the divisibility relation.

Example: $X \subset \mathbb{Z}$ with Divisibility. Now consider the set X consisting of integers expressible as a product of at most 5 prime factors. For example, X contains $2^3 3^1 = 24$ and $2^3 3^1 7^1 = 168$ but *not* $2^3 3^1 5^1 7^1 = 840$. We use divisibility as our relation on X. Every chain in X has at most six elements. Moreover every chain $C \subset X$ has an upper bound: either $C = \emptyset$, in which case 1 (or any element of X) is an upper bound for C, or the largest element of C is an upper bound for C. The element $32 \in X$ (or, for that matter, *any* element with exactly 5 prime factors, not necessarily distinct) is a maximal element of X. Note, however, that 32 is *not* an upper bound for X.

Zorn's Lemma. Let X be a nonempty partially ordered set, and suppose every chain in X is bounded above. Then X has a maximal element.

Like most authors, we assume this result rather than proving it. The reason for this is that one cannot prove this result without assuming the Axiom of Choice (or something at least as strong). This is because Zorn's Lemma is equivalent to the Axiom of Choice, given the Zermelo-Fraenkel axioms of set theory. It is typically used as a convenient crutch, where no maximal element is explicitly constructible. This should not be of great concern, however, since in practical situations where a maximal element is desired, we can typically get by without one. We will try to make this point clear in the context of an example.

Corollary. Every vector space has a basis.

Proof. Let V be a vector space over a field F. We assume $V \neq 0$; otherwise $\emptyset$ is a basis for V.

Let $\mathcal{I}$ be the collection of all linearly independent subsets of V. Recall that a subset $S \subseteq V$ is *linearly dependent* if there exist distinct vectors $v_1, v_2, \ldots, v_k \in S$ and scalars $a_1, a_2, \ldots, a_k \in F$, not all zero, such that $a_1v_1 + a_2v_2 + \cdots + a_kv_k = 0$. Thus $S \in \mathcal{I}$ iff $S \subset V$ and S is *not* linearly dependent. Clearly $\mathcal{I}$ is nonempty, since every nonzero vector $v \in V$ gives rise to a linearly independent subset $\{v\} \in \mathcal{I}$.

Let $\mathcal{C} \subset I$ be any chain. We claim that $\mathcal{C}$ is bounded above by $\bigcup \mathcal{C}$. (Recall that $\bigcup \mathcal{C}$ is the union of all members of $\mathcal{C}$; that is, $\bigcup \mathcal{C} = \bigcup_{S \in \mathcal{C}} S$.) We must first show that $\bigcup \mathcal{C} \in \mathcal{I}$. Consider any distinct vectors $v_1, v_2, \ldots, v_k \in \bigcup \mathcal{C}$ and let $a_1, a_2, \ldots, a_k \in F$. For every $i = 1, 2, \ldots, k$, the fact that $v_i \in \bigcup \mathcal{C}$ means that $v_i \in S_i$ for some linearly independent subset $S_i \in \mathcal{C}$. Since $\mathcal{C}$ is a chain, the S_i's are totally ordered by inclusion. This means we may assume that $S_1 \subseteq S_2 \subseteq \cdots \subseteq S_k$; at least this will be the case if $v_1, v_2, \ldots, v_k$ were listed in a suitable order. But now $v_1, v_2, \ldots, v_k$ all belong to the linearly independent set S_k, and so the scalars $a_1, a_2, \ldots, a_k$ must all be zero. This shows that $\mathcal{C}$ is linearly independent, so $\bigcup \mathcal{C} \in \mathcal{I}$. We still need to show that $\bigcup \mathcal{C}$ is an upper bound for the chain $\mathcal{C}$. But this is obvious since for every linearly independent subset $S \in \mathcal{C}$, we have $S \subseteq \bigcup \mathcal{C}$ by definition.

Let B be a maximal element for $\mathcal{I}$, which exists by Zorn's Lemma. So B is linearly independent. It remains to be shown that B spans V. Let $v \in V$. We must show that v is in the span of B. If $v \in B$ then this is clear; so we may assume that $v \notin B$, so that B is a proper subset of $B \cup \{v\}$. Since B is a maximal element of $\mathcal{I}$, it must be the case that $B \cup \{v\}$ is linearly dependent. Thus there exist distinct vectors $v_1, v_2, \ldots, v_k \in B \cup \{v\}$ and scalars $a_1, a_2, \ldots, a_k \in F$, not all zero, such that

$$a_1v_1 + a_2v_2 + \cdots + a_kv_k = 0.$$

Clearly $v \in \{v_1, v_2, \ldots, v_k\}$ since B itself is linearly independent; we may assume that $v_1 = v$. Moreover $a_1 \neq 0$, for otherwise we have found a nontrivial linear relation between $v_2, v_3, \ldots, v_k \in B$, which cannot occur since B is linearly dependent. Thus

$$v = -a_1^{-1}\big(a_2v_2 + a_3v_3 + \cdots + a_kv_k\big)$$

lies in the span of B, as required. Thus B spans V. Since B is also linearly independent, B is a basis for V. $\square$

For finite dimensional vector spaces, it is very easy to produce bases explicitly, and so Zorn's Lemma is not needed in such cases. For many infinite-dimensional vector spaces, this is not an option. For example, the vector space $C([0,1])$ consisting of continuous functions $[0,1] \to \mathbb{R}$, has a basis, by Zorn's Lemma. But you will never see an explicit basis for this vector space! since none can be written down. But in any practical situation in which $C([0,1])$ arises, this is not an issue since we typically deal with only certain well-known proper subspaces of $C([0,1])$ for which explicit bases are known.