Zorn's lemma and some examples of its application

Zorn's lemma is an extremely handy tool for dealing with constructions that require infinitely many steps to be made. Consider the following theorem:

Theorem 1. *Let R be a unital ring and M be an R-module. Then M has a maximal linearly independent subset (that is, a linearly independent subset that is not contained in any other linearly independent subset of M).*

One can try to naively prove this theorem this way: "If $\emptyset$ is not the maximal linearly independent subset of M, choose $u_1 \in M$ such that the set $\{u_1\}$ is linearly indepedent. If $\{u_1\}$ is not a maximal linearly independent subset of M, choose $u_2 \in M$ such that the set $\{u_1, u_2\}$ is linearly indepedent. And so on, until we get a maximal linearly independent subset of M." This argument does not look convincing because the process described above may never end, – even after infinitely, uncountably many steps, – so that we will never get the maximal linearly independent subset of M. Zorn's lemma is designed to convert arguments like this into rigorous proofs. Here is its standard formulation:

Zorn's lemma. *Let S be a partially ordered set in which every chain has an upper bound. Then S has a maximal element.*

(Terminology: A set S is partially ordered if *a partial order* "$<$" is introduced on it, that is, for some pairs $a, b \in S$ one has $a < b$ so that it can never be that both $a < b$ and $b < a$, and so that $a < b < c$ implies $a < c$. *A chain* in S is a totally ordered subset $C \subseteq S$, which means that any two elements of C are comparable: for any distinct $a, b \in C$, either $a < b$ or $b < a$. *An upper bound* of a set $C \subseteq S$ is an element $c \in S$ such that $a \leq c$ for all $a \in C$. An element $a \in S$ is *maximal* if there is no $b \in S$ such that $a < b$.)

In most cases the following special form of Zorn's lemma applies:

Zorn's lemma 2. *Let X be a set and let $\mathcal{S}$ be a family of subsets of X such that for any chain $\mathcal{C}$ in $\mathcal{S}$ one has $\bigcup_{A \in \mathcal{C}} A \in \mathcal{S}$. Then $\mathcal{S}$ has a maximal element.*

(In this formulation, the order on $\mathcal{S}$ is given by the (strict) inclusion "$\subsetneqq$" relation, so that a chain in $\mathcal{S}$ is a subfamily $\mathcal{C} \subseteq \mathcal{S}$ such that for any distinct $A, B \in \mathcal{C}$ either $A \subset B$ or $B \subset A$, and A is a maximal element of $\mathcal{S}$ if there is no $B \in \mathcal{S}$ such that $A \subsetneqq B$.)

We may now prove Theorem 1:

Proof of Theorem 1. Let $\mathcal{S}$ be the family of all linearly independent subsets of M. If $\mathcal{C}$ is a chain in $\mathcal{S}$, then the set $D = \bigcup_{A \in \mathcal{C}} A$ is linearly independent. (Indeed, for any $u_1, \ldots, u_n \in D$ we have $u_i \in A_i$ for some $A_i \in \mathcal{C}$, $i = 1, \ldots, n$, and since all A_i are comparable, one of them, say A_k, contains all others; so $u_1, \ldots, u_n \in A_k$ and so, $u_1, \ldots, u_n$ are linearly independent.) Thus, $D \in \mathcal{S}$. Hence, Zorn's lemma applies to $\mathcal{S}$ and guarantees that there is a maximal linearly independent set B in M. ■

Here is another example:

Theorem 2. *Let R be a commutative unital ring; then R has a maximal proper ideal.*

Proof. Let $\mathcal{S}$ be the set of all proper ideals in R. For any chain $\mathcal{C} \subseteq \mathcal{S}$, $J = \bigcup_{I \in \mathcal{C}} I \in \mathcal{S}$. (Indeed, if $a, b \in J$, then $a \in I_1$, $b \in I_2$ for some $I_1, I_2 \in \mathcal{C}$. Assume, w.l.o.g., that $I_1 \subseteq I_2$; then $a_1, a_2 \in I_2$, so $r_1 a_1 + r_2 a_2 \in I_2 \subseteq J$ for any $r_1, r_2 \in R$. So, J is an ideal. Also, $J \neq R$ since $1 \notin I$ for all $I \in C$, so $1 \notin J$.) So, Zorn's lemma applies, and says that R has a maximal ideal. ■

The next example, called *the replacement lemma*, is more sophisticated:

Theorem 3. *Let V be a vector space over a field F, let A be a linearly independent subset of V, and let B be a spanning subset of V. Then $|B| \geq |A|$, where $|A|$ and $|B|$ are the cardinalities of A and B respectively. In particular, any two bases in V have the same cardinality (called the dimension of V).*

Proof. Let $\mathcal{S}$ be the set of triplets $\tau = (P, Q, \varphi)$ where $P \subseteq A$, $Q \subseteq B$, $P \cup Q$ is linearly independent, and φ is a bijection $Q \longrightarrow (A \setminus P)$; $\mathcal{S}$ is nonempty since $(A, \emptyset, \phi) \in \mathcal{S}$ where $\phi\colon \emptyset \longrightarrow (A \setminus A)$. A partial order on $\mathcal{S}$ is introduced in the following way: we have $(P_1, Q_1, \varphi_1) < (P_2, Q_2, \varphi_2)$ if $P_2 \subset P_1$, $Q_1 \subset Q_2$, and $\varphi_2|_{Q_1} = \varphi_1$. For $\tau = (P, Q, \varphi) \in \mathcal{S}$, define $P_\tau = P$, $Q_\tau = Q$, $\varphi_\tau = \varphi$.

Now, given a chain $\mathcal{C}$ in $\mathcal{S}$, put $\widehat{P} = \bigcap_{\tau \in \mathcal{C}} P_\tau$, $\widehat{Q} = \bigcup_{\tau \in \mathcal{C}} Q_\tau$, and $\widehat{\varphi} = \bigcup_{\tau \in \mathcal{C}} \varphi_\tau$. (That is, define $\widehat{\varphi}\colon \widehat{Q} \longrightarrow (A \setminus \widehat{P})$ by $\widehat{\varphi}(u) = \varphi_\tau(u)$ for any $\tau \in \mathcal{S}$ for which $u \in Q_\tau$; since φ_{τ_2} is an extension of φ_{τ_1} if $\tau_2 > \tau_1$, $\widehat{\varphi}(u)$ does not depend on the choice of $\tau \in \mathcal{C}$.) $(\widehat{P}, \widehat{Q}, \widehat{\varphi})$ is an upper bound of $\mathcal{C}$, as soon as we show that it is an element of $\mathcal{S}$.

Clearly, $\widehat{\varphi}$ is a bijection: every two elements of $\widehat{Q}$ belong to Q_τ for some $\tau \in \mathcal{C}$, and so, have distinct images under $\widehat{\varphi}$; and every element of $A \setminus \widehat{P}$ is contained in $A \setminus P_\tau$ for some $\tau \in \mathcal{C}$, and hence, is an image of some element of $Q_\tau \subseteq \widehat{Q}$. To prove that $(\widehat{P}, \widehat{Q}, \widehat{\varphi}) \in \mathcal{S}$, we only need to show that the set $\widehat{P} \cup \widehat{Q}$ is linearly independent. Indeed, any linear combination of elements of $\widehat{P} \cup \widehat{Q}$ only involves finitely many elements of $\widehat{Q}$, all these elements are contained in Q_τ for some $\tau \in \mathcal{C}$, so the combination only involves elements of $P_\tau \cup Q_\tau$, which is a linearly independent set; hence, $\widehat{P} \cup \widehat{Q}$ is linearly independent.

Since every chain in $\mathcal{S}$ has an upper bound, Zorn's lemma applies and says that $\mathcal{S}$ has a maximal element (P, Q, φ). Assume that $P \neq \emptyset$, let $u \in P$. Put $P' = P \setminus \{u\}$. Since B spans V, we have $u = a_1 v_1 + \cdots + a_k v_k$ for some $v_1, \ldots, v_k \in B$ and $a_1, \ldots, a_k \in F \setminus \{0\}$. Since $\{u\} \cup P' \cup Q$ is linearly independent, at least one of $v_i \notin \mathrm{Span}(P' \cup Q)$; w.l.o.g. assume that $v_1 \notin \mathrm{Span}(P' \cup Q)$. Put $Q' = Q \cup \{v_1\}$, then $P' \cup Q'$ is linearly independent. Consider the triplet (P', Q', φ') where $\varphi'\colon Q' \longrightarrow (A \setminus P')$ is defined by $\varphi'|_Q = \varphi$ and $\varphi'(v_1) = u$. Then $(P', Q', \varphi') \in \mathcal{S}$ and $(P', Q', \varphi') > (P, Q, \varphi)$, which contradicts the maximality of (P, Q, φ). Hence, $P = \emptyset$. But then φ is a bijection between $Q \subseteq B$ and $A \setminus \emptyset = A$, so $|A| \leq |B|$. ■