This is a short handout discussing Zorn's Lemma. Another good reference for this material (from a Fields medalist) is located here.

Let's recall the definition of a partial order:

Definition 0.1. A *partial order* on a set A is a (binary) relation $\preceq$ that is reflexive ($a \preceq a$ for all $a \in A$), antisymmetric (if $a \preceq b$ and $b \preceq a$, then $a = b$), and transitive (if $a \preceq b$ and $b \preceq c$, then $a \preceq c$). If A has a partial ordering $\preceq$, we call $(A, \preceq)$ a *poset* (*p*artially *o*rdered *set*). If the partial order $\preceq$ on A is clear from context, we just write A instead.

Note that the definition of partial order does not require any two elements a and b in A to be comparable, by which we mean either $a \preceq b$ or $b \preceq a$ (for instance, consider the subset $\subseteq$ partial order on the power set of the set $\{1, 2, 3\}$). But if this is true, we give a special name:

Definition 0.2. A partial order on A in which any two elements are comparable (i.e. either $a \preceq b$ or $b \preceq a$) is called a *total order*. We then call A *totally ordered*.

For instance, the usual "less than or equal to" ordering $\leq$ on $\mathbf{R}$ is a total order.

If A' is a subset of A, and $\preceq$ is a partial order on A, then we get a partial order on A' simply by restricting $\preceq$ to A'. Now, $\preceq$ may not be a total order on A, but it may become one once restricted to a subset A'. This situation is important, so we give it a name:

Definition 0.3. Let A be a poset, and let A' be a nonempty subset of A such that the partial ordering $\preceq$ becomes a total order when restricted to A' (that is, for every $a'_1, a'_2 \in A'$, either $a'_1 \preceq a'_2$ or $a'_2 \preceq a'_1$). We then call A' a *chain* in A.

Of course, when A is itself totally ordered, then any subset is a chain. So this definition is only nontrivial when A is not totally ordered.

Example 0.1. For instance, $\mathcal{P}(\mathbf{N})$ is a poset under $\subseteq$, and is not totally ordered. But the subset $\{[n] : n \in \mathbf{N}\} \subseteq \mathcal{P}(\mathbf{N})$ is a chain.

Note that a chain *need not be countable*. As an example, consider $\mathcal{P}(\mathbf{R})$, which is a non-totally ordered poset under $\subseteq$. For any real number r, let $C_r = \{s \in \mathbf{R} : s < r\}$. If r, r' are distinct real numbers, then $C_r \neq C_{r'}$, so there are uncountably many C_r. The subset of C_r's inside $\mathcal{P}(\mathbf{R})$ forms an uncountable chain.

We now give two definitions relating to maximality, before we can state Zorn's Lemma.

Definition 0.4. Let A be a poset and let S be a subset of A (not necessarily a chain). We say $a \in A$ is an *upper bound* for S if, for all $s \in S$, we have $s \preceq a$. Note that the definition implicitly requires a to be comparable to every element in S. On the other hand, a need not be in S.

Example 0.2. Consider $\mathbf{R}$ with the usual ordering $\leq$, which turns it into a poset (even totally ordered). Then an upper bound for $(0, 1)$ is 1. On the other hand, the subset $\mathbf{Q}$ has no upper bound.

Example 0.3. Let A be a nonempty set, and let $\mathcal{P}(A)$ be the power set of A with the partial ordering $\subseteq$. Then $A \in \mathcal{P}(A)$ itself is an upper bound for any subset of $\mathcal{P}(A)$.

Definition 0.5. Let A be a poset. Then a *maximal element* of A is an element $a \in A$ such that for any $b \in A$ comparable to a has $b \preceq a$. In particular, we do not require that a be comparable to every element in A.

Note that a poset A can have *any number* of maximal elements, including infinitely many. As examples, $\mathbf{R}$ with the usual order has no maximal element. $\mathcal{P}(\{1, 2, 3\})$ (with the usual partial ordering) has one maximal element, $\{1, 2, 3\}$, but the subset of $\mathcal{P}(\{1, 2, 3\})$ not containing $\{1, 2, 3\}$ has *three* maximal elements: $\{1, 2\}$, $\{1, 3\}$, and $\{2, 3\}$. Similarly, $\mathcal{P}(\mathbf{R})$ has a single maximal element, but the subset of $\mathcal{P}(\mathbf{R})$ not containing $\mathbf{R}$ has (uncountably) infinitely many maximal elements (what are they?).

Here are some basic facts about maximal elements, which are good to prove as exercises.

Exercise 0.1. If A is a totally ordered set, then A has at most one maximal element.

Exercise 0.2. If A is a totally ordered set, then any finite subset of A has exactly one maximal element.

Exercise 0.3. Suppose A is a poset, and $a \in A$ is a maximal element. If $a \preceq b$ for some $b \in A$, then $a = b$.

We can now state Zorn's Lemma:

Lemma 0.1 (Zorn)**.** Let A be a nonempty poset such that every (nonempty) chain in A has an upper bound. Then A has at least one maximal element.

We will not prove Zorn's Lemma, but only mention that it is equivalent to a foundational *axiom* of set theory: the axiom of choice. The axiom of choice is a statement that seems incredibly obvious at first, but turns out to be *independent* from the rest of the axioms of set theory (in the sense that it cannot be proved from the other axioms), and can lead to many surprising and unintuitive consequences (Zorn's lemma, and then Theorem 0.2, for instance).

Intuitively, what Zorn's lemma (and its equivalent, the axiom of choice) allows you to do is to "make uncountably many choices at once." For instance, we often might try to give a recursive definition/construction of some object, where we build our object in "stages," making some choice at each stage. If the number of choices we need to make is finite or countably infinite, we should be more or less fine, because either our process stops, or we

just keep building one stage at a time. But if we need to make *uncountably many choices*, then we have a problem, because we'll never manage to make all the choices we need (making one choice at a time implies some sort of countability in our process). Here is a quote from the Fields medalist Timothy Gowers that nicely summarizes the utility of Zorn's Lemma:

> If you are building a mathematical object in stages and find that (i) you have not finished even after infinitely many stages, and (ii) there seems to be nothing to stop you continuing to build, then Zorn's lemma may well be able to help you.

Zorn's Lemma is used in many fundamental theorems in all areas of math, not just set theory. For example, the following is a small sample of theorems, all of which are proved using Zorn's Lemma:

- Every ring contains a maximal ideal (abstract algebra).
- Every field has an algebraic closure (abstract algebra).
- Tychonoff's theorem (topology).
- Hahn-Banach theorem (functional analysis).

In each of these cases, the objects of interest are equipped with some additional structure (algebraic, topological, analytic), and these structures need to be remembered during the steps of the proof, but Zorn's lemma is what allows one to make the ultimate construction.

We now give a sample application of Zorn's Lemma in linear algebra.

Theorem 0.1. Let F be a field (if you don't know what a field is, replace F with the real numbers $\mathbf{R}$ throughout). Then every F-vector space V has a basis.

Note that the basis—a linearly independent spanning set—can be infinite. For a possibly infinite set of vectors $\{v_i\}$ in V, we say that it is a spanning set for V if any $v \in V$ is a *finite* F-linear combination of some of the v_i's, and any *finite* subset of vectors in $\{v_i\}$ is linearly independent in the usual sense. Of course, when the set $\{v_i\}$ is finite, this is just the usual definition from finite-dimensional linear algebra.

In previous linear algebra classes, you probably saw and/or proved this theorem when V was *finite dimensional*. But here, we *do not* make that assumption! Therefore any proof you may have encountered previously, which probably involved taking a *finite* spanning set and throwing out any "redundant" elements of that spanning set, does not work in this case, because here a spanning set for V over F may even be uncountably infinite, so our "throwing out" process may never end. Similarly, any proof that involved "building a basis by adding one vector at a time" cannot work, again because our "adding" process may need to involve uncountably many choices, and therefore never end. For example, $\mathbf{R}$ is a vector space over

$\mathbf{Q}$, but any spanning set (hence any basis) of $\mathbf{R}$ over $\mathbf{Q}$ must be *uncountable* for cardinality reasons, so it is *impossible* to even explicitly write down any such spanning set.

We will now prove Theorem 0.1 (it looks long, but that's more due to writing out all the details—once you get more practice with Zorn's Lemma, the main idea will seem quite straightforward). The idea is to mimic the proof of "adding one element at a time", but using Zorn's lemma to make uncountably many choices all at once!

Proof. First, we have the trivial case when V is the zero vector space $\{0\}$. Then the empty set is a basis for V (by convention, or by some post-hoc reasoning with empty set nonsense), so this case is proved. If this doesn't sit right with you, feel free to ignore it and just add the additional hypothesis that V is a nontrivial vector space in the statement of the theorem.

Now suppose V is nontrivial. Let S be the set of all (nonempty) linearly independent subsets of V. Since V is nontrivial, it has some nonzero vector v, so $\{v\}$ is a linearly independent subset of V. Hence S is nonempty. We can give a partial ordering on S by the usual $\subseteq$ inclusion ordering.

Our intuition from the finite-dimensional case tells us that a basis should be a "maximal linearly independent subset." Zorn's Lemma tells us how to prove that one exists! So, we need to take an arbitrary chain $C = \{C_i\}$ in S, so C is a totally ordered subset of S, and each element C_i of C is a linearly independent subset of V. Let $\{v_j\}$ be the *union* of all elements in C (this is the common trick when applying Zorn's Lemma: take the union of the chain). We need to show that $\{v_j\}$ is also a linearly independent subset. Indeed, let $\{v_1, \ldots, v_n\}$ be a finite subset of $\{v_j\}$, and suppose there is a linear combination

$$a_1 v_1 + a_2 v_2 + \ldots + a_n v_n = 0,$$

with the $a_i \in F$. Then each v_k, for $1 \leq k \leq n$, comes from some C_{i_k}. But the C_i's are *totally ordered* by definition of a chain, so there is some maximal element in the *finite subset* $\{C_{i_1}, C_{i_2}, \ldots, C_{i_n}\}$. Since our partial ordering is inclusion, this means one of the C_{i_k}'s, say C_{i_l}, contains all the rest of the C_{i_k}'s. Then the $v_1, \ldots, v_n$ are all vectors in C_{i_l}, but C_{i_l} is a *linearly independent subset* of V by definition. Therefore $a_1 = a_2 = \ldots = a_n = 0$, and since $\{v_1, \ldots, v_n\}$ was an arbitrary finite subset of $\{v_j\}$, we conclude that $\{v_j\}$ is a linearly independent subset of V by definition. Hence it is an element of S.

So $\{v_j\}$, which is in S, is visibly an upper bound for the chain C (even though it might not be in C). Since C was an arbitrary chain, we satisfy the hypotheses of Zorn's Lemma, and so we produce a maximal element $\{b_i\}$ of S. Again, our intuition tells us that $\{b_i\}$ is a basis for V. It is linearly independent by definition of S, so we just need to check that it spans V.

Assume for the sake of contradiction that there is some $v \in V$ not in the span of $\{b_i\}$ (so in particular v is not equal to any of the b_i). Then we claim that $\{b_i\} \cup \{v\}$ is a linearly independent subset of V. Suppose there is a linear combination

$$a_1 b_1 + \ldots + a_n b_n + a_v v = 0.$$

Then a_v cannot be nonzero, as otherwise

$$v = -\frac{1}{a_v}\left(a_1 b_1 + \ldots + a_n b_n\right),$$

which would write v as a linear combination of the vectors in $\{v_i\}$, contrary to the definition of v. Then $a_v = 0$, and since we know that $\{b_i\}$ is a set of linearly independent vectors, we must have $a_1 = \ldots = a_n = 0$. Therefore $\{b_i\} \cup \{v\}$ is a linearly independent subset of V, so in S. But this contradicts the *maximality*(!!) of $\{b_i\}$: we have an element $\{b_i\} \cup \{v\}$ that is comparable to $\{b_i\}$, but $\{b_i\} \cup \{v\} \not\subseteq \{b_i\}$. Therefore there can be no $v \in V$ that is not in the span of $\{b_i\}$, so $\{b_i\}$ is a linearly independent spanning set of V, hence a basis.

Notice that this last part of the proof is the exact same as in the finite-dimensional case. □

To internalize the above argument, it's instructive to locate where and how each hypothesis (and the conclusion) of Zorn's Lemma was used in the above proof.

Here is another accessible application of Zorn's Lemma, which is a purely set-theoretical statement:

Theorem 0.2 (Well-ordering)**.** Let A be a nonempty set. Then there *exists* a total order on A that is a well-ordering (i.e. an order where every nonempty subset $S \subseteq A$ has a least element: an element $s \in S$ such that $s \preceq s'$ for all $s' \in S$).

Proof. The proof can be found here—it is a piece of somewhat technical set-theoretic reasoning. □

Notice that we do not say that *any* total ordering on A is a well-ordering. For instance, when $A = \mathbf{R}$, the usual ordering given by $\leq$ is not a well-ordering, because the subset $(0, 1)$ does not have a least element. In fact, it is *impossible* to explicitly write down (in any reasonable sense) a well-ordering on $\mathbf{R}$, because we would have to be making uncountably many choices at once (precisely what Zorn's Lemma helps us do!). Zorn's Lemma can only tell us that one *exists*.