Existence of Bases of an Arbitrary Vector Space

Mathematics 481/525

In most textbooks authors prove that every finitely-generated vector space has a basis. In this note we will eliminate the finitely generated hypothesis. We first define what this means. Let $F$ be a field and $V$ an $F$-vector space. A subset $S$ of $V$ is linearly independent if any finite subset $\{v_1, \ldots, v_n\}$ of $S$ is linearly independent. In other words, $S$ is linearly independent if for each $n \in \mathbb{N}$ and $v_1, \ldots, v_n \in S$, the only solution to $\sum_{i=1}^{n} x_i v_i = 0$ is $x_1 = \cdots = x_n = 0$. Recall that $S$ spans $V$ if $V$ is the only subspace which contains $S$. What this means is that each $v \in V$ is a linear combination of finitely many elements of $S$. A basis of $V$ is a subset which is both linearly independent and spans $V$. A nice characterization is that a subset $S$ of $V$ is a basis of $V$ if every element of $V$ is uniquely expressible as a linear combination of finitely many elements of $V$.

Example 1. Let $V = F[x]$. Then $\{x^n : n \geq 0\}$ is a basis of $V$, since every $f(x) \in V$ is uniquely expressible as a linear combination of the monomials $\{x^n\}$. For example, $1+3x+5x^4$ is a linear combination of $x^0, x, x^4$.

To prove the existence of a basis for an arbitrary vector space, we need more sophisticated tools than we have been using. We will use is a statement equivalent to the axiom of choice. While we wont’ use it, the axiom of choice states that if $\{A_i : i \in I\}$ is a family of nonempty sets, then there is a function $f : I \rightarrow \bigcup_{i \in I} A_i$ for which $f(i) \in A_i$ for each $i$. The meaning of the name is that we are “choosing” an element $a_i$ from each $A_i$, and defining $f$ by $f(i) = a_i$. The axiom of choice is named thusly because it is an axiom of set theory.

In this note we will use Zorn’s lemma, a theorem equivalent to the axiom of choice. To state it, we need some preliminary notations. If $X$ is a set, then a relation $\leq$ on $X$ is said to be a partial order if the relation is reflexive, transitive, and anti-symmetric; this latter condition means if $a \leq b$ and $b \leq a$, then $a = b$. We will refer to the pair $(X, \leq)$ as a partially ordered set, or, for short, a poset.

Example 2. Consider the power set of a set $T$. By defining $A \leq B$ if $A$ and $B$ are subsets of $T$ with $A \subseteq B$, then the power set is a poset with this relation. It is quite possible to have subsets $A, B$ with $A \not\subseteq B$ and $B \not\subseteq A$.

Example 3. The set $\mathbb{R}$ with its usual order is a poset. Note that if $a, b \in \mathbb{R}$, then either $a \leq b$ or $b \leq a$. 
If \((X, \leq)\) is a poset and \(a, b \in X\), we need not have either \(a \leq b\) or \(b \leq a\). However, if for each pair \(a, b \in X\), either \(a \leq b\) or \(b \leq a\), then we call \((X, \leq)\) a chain. If \(C\) is a subset of \(X\) which, when restricting the partial order to \(C\), is a chain, then we say \(C\) is a chain in \(X\). The set of real numbers with its usual ordering is a chain.

Let \((X, \leq)\) be a poset. If \(A\) is a subset of \(X\), then \(x \in X\) is said to be an upper bound for \(A\) if \(a \leq x\) for each \(a \in A\). Note that we are not assuming that \(x \in A\). A maximal element of \(A\) is an element \(b \in A\) such that if \(a \in A\) with \(b \leq a\), then \(a = b\). In other words, there is no properly larger element in \(A\) than \(b\). This does not say that \(a \leq b\) for all \(a \in A\) (such an element would be called a maximum element).

**Example 4.** Let \(X\) be the power set of \(\{1, 2, 3, 4\}\) with the partial order of inclusion. If \(A = \{\emptyset, \{1, 2\}, \{2, 3\}, \{1, 2, 4\}\}\), then only \(\{1, 2, 3, 4\}\) is an upper bound for \(A\). However, \(\{2, 3\}\) and \(\{1, 2, 4\}\) are maximal elements of \(A\). The subposet \(\{\emptyset, \{1\}, \{1, 3\}\}\) is a chain in \(X\).

We can now state Zorn’s lemma.

**Theorem 5 (Zorn’s Lemma).** Let \((X, \leq)\) be a nonempty poset. If each chain \(C\) in \(X\) has an upper bound in \(X\), then \(X\) has a maximal element.

To apply Zorn’s lemma we only need to know that each chain has an upper bound in \(X\); the upper bound need not be in the chain. With Zorn’s lemma, we will prove the existence of maximal ideals in rings with 1 and the existence of bases of vector spaces. The way we apply Zorn’s lemma in this note are typical applications of this result in algebra.

**Theorem 6.** Let \(R \neq \{0\}\) be a ring with 1. Then \(R\) has a maximal ideal.

**Proof.** Let \(S\) be the set of all proper ideals of \(R\). Then \(S\) is nonempty since \((0) \in S\). We partially order \(S\) by inclusion. In other words, we define \(I \leq J\) if \(I \subseteq J\). It is trivial to check that \((S, \subseteq)\) is a poset. To apply Zorn’s lemma, let \(C = \{I_\alpha : \alpha \in \Gamma\}\) be a chain in \(S\), indexed by some set \(\Gamma\). We must prove that \(C\) has an upper bound in \(S\). Let \(I = \bigcup_{\alpha \in \Gamma} I_\alpha\), the union of all the \(I_\alpha\). We claim that \(I\) is a proper ideal of \(R\). To see that \(I\) is an ideal, first we show that \(I\) is a subgroup of \((R, +)\). Let \(a, b \in I\). Then there are \(\alpha, \beta \in \Gamma\) with \(a \in I_\alpha\) and \(b \in I_\beta\). Since \(C\) is a chain, either \(I_\alpha \subseteq I_\beta\) or \(I_\beta \subseteq I_\alpha\); without loss of generality we assume that \(I_\alpha \subseteq I_\beta\). Then \(a, b \in I_\beta\). Since \(I_\beta\) is an ideal of \(R\), we have \(a - b \in I_\beta \subseteq I\). Therefore, \(I\) is a subgroup of \((R, +)\). Next, let \(a \in I\) and \(r \in R\). Then \(a \in I_\alpha\) for some \(\alpha\). Since \(I_\alpha\) is an ideal of \(R\), both \(ra\) and \(ar\) are elements of \(I_\alpha\). Thus, both are elements of \(I\). This proves that \(I\) is a proper ideal of \(R\). To see that it is a proper ideal, note that \(1 \notin I_\alpha\) for each \(\alpha\) since each \(I_\alpha \in S\) is a proper ideal of \(R\). Therefore, \(1 \notin I\) by definition of \(I\), so \(I\) is indeed proper. It is an upper bound for the chain since \(I_\alpha \subseteq I\) for each \(\alpha \in \Gamma\). So, the hypotheses of Zorn’s lemma applies. There is then a maximal element \(M\) of \(S\). We claim that \(M\) is a maximal ideal of \(R\). To prove this, suppose that \(M \subseteq J \subseteq R\) with \(J\) an ideal of \(R\). If \(J\) is proper, then \(J \in S\). Maximality of \(M\) then forces \(M = J\). Thus, either \(J = M\) or \(J = R\).
We next prove the existence of a basis for each vector space. By carefully thinking about the definitions, we see that \( \emptyset \) is a basis of the zero vector space \( \{0\} \); more generally, \( \emptyset \) is a linearly independent subset of any vector space. To prove the theorem, we need the following lemma.

**Lemma 7.** Let \( V \) be a vector space. If \( B \) is a linearly independent subset of \( V \), and if \( v \in V \) is not in the span of \( B \), then \( B \cup \{v\} \) is also linearly independent.

**Proof.** To prove that \( B \cup \{v\} \) is linearly independent, we must show that each finite subset is linearly independent. Take \( \{v_1, \ldots, v_n\} \subseteq B \cup \{v\} \). We have two cases. First, if \( v \notin \{v_1, \ldots, v_n\} \), then \( \{v_1, \ldots, v_n\} \subseteq B \), so it is linearly independent since \( B \) is linearly independent. If, on the other hand, \( v \in \{v_1, \ldots, v_n\} \), then, after changing notation if necessary, we set \( v_1 = v \). Suppose that \( \sum_{i=1}^n \alpha_i v_i = 0 \). If \( \alpha_1 \neq 0 \), then \( v_1 = \sum_{i=2}^n -\alpha_i \alpha_1^{-1} v_i \), which says \( v = v_1 \) is in the span of \( B \). Since this is false, \( \alpha_1 = 0 \). Then \( \sum_{i=2}^n \alpha_i v_i = 0 \), which by the linear independence of \( B \), noting that \( \{v_2, \ldots, v_n\} \subseteq B \), implies that \( \alpha_2 = \cdots = \alpha_n \). Thus, all \( \alpha_i = 0 \), proving that \( \{v_1, \ldots, v_n\} \) is linearly independent. \( \square \)

**Theorem 8.** Every vector space has a basis.

**Proof.** Let \( V \) be an \( F \)-vector space, and let \( S \) be the set of linearly independent subsets of \( V \). The set \( S \) is nonempty since \( \emptyset \in S \). We partially order \( S \) by inclusion. Let \( C = \{A_\alpha : \alpha \in \Gamma\} \) be a chain in \( S \), and set \( A = \bigcup_{\alpha \in \Gamma} A_\alpha \). We claim that \( A \) is an upper bound in \( S \) of the chain \( C \). We must prove that \( A \) is linearly independent; that it is an upper bound of \( C \) then follows immediately by the definition of \( A \). So, suppose that \( v_1, \ldots, v_n \in A \) and that \( \sum_{i=1}^n b_i v_i = 0 \) for some \( b_i \in F \). Each \( v_i \) is an element of some \( A_{\alpha_i} \in C \). Because \( C \) is a chain, there is some \( A_{\alpha_j} \) containing each \( A_{\alpha_1}, \ldots, A_{\alpha_n} \). Then \( \{v_1, \ldots, v_n\} \) is a subset of \( A_{\alpha_j} \), a linearly independent subset of \( V \). This forces \( b_1 = \cdots = b_n = 0 \). Therefore, \( A \) is indeed linearly independent. So, by Zorn’s lemma, \( S \) has a maximal element \( B \). We claim that \( B \) is a basis for \( V \). It is linearly independent by the fact that \( B \in S \). To prove \( B \) spans \( V \), let \( v \in V \). If \( v \) is not in the span of \( B \), then \( B \cup \{v\} \) is linearly independent by the lemma. However, this would contradict the maximality of \( B \). Therefore, \( v \) is in the span of \( B \). This proves that \( V \) is the span of \( B \), and so \( B \) is a basis of \( V \). \( \square \)

We can modify the argument above to get some extensions of this theorem.

**Proposition 9.** Let \( V \) be an \( F \)-vector space.

(1) Let \( A \) be a linearly independent subset of \( V \). Then \( A \) is contained in a basis of \( V \).

(2) Let \( S \) be a spanning set of \( V \). Then \( S \) contains a basis of \( V \).

(3) Let \( A \subseteq S \) be subsets of \( V \) such that \( A \) is linearly independent and \( S \) spans \( V \). Then there is a basis \( B \) of \( V \) with \( A \subseteq B \subseteq S \).
Proof. We will prove (3), the strongest of the three. Statement (1) follows from (3) by setting $S = V$, and (2) follows from (3) by setting $A = \emptyset$. To prove (3), let

$$S = \{ D : A \subseteq D \subseteq S, \text{ } D \text{ is linearly independent} \}.$$  

Then $S$ is nonempty since $A \in S$. Just as in the proof of the previous theorem, Zorn’s lemma applies to $S$. Therefore, there is a maximal element $B$ of $S$. Then $A \subseteq B \subseteq S$ by definition of $S$. We claim that $B$ is a basis of $V$. That $B$ is linearly independent follows from $B \in S$. To see that $B$ spans $V$, take $v \in V$. Then $v = \alpha_1 w_1 + \cdots + \alpha_n w_n$ for some $\alpha_i \in F$ and $w_i \in S$; we are using that $S$ spans $V$ to do this. For each $i$ we see that $w_i$ is in the span of $B$, since by the lemma $B \cup \{ w_i \}$ would be linearly independent, which is false since then $B \cup \{ w_i \} \in S$. Thus, each $w_i$ is in the span of $B$. Therefore, $v$ is also in the span of $B$. This proves that $B$ spans $V$. 

Now that we know every vector space has a basis, we can ask the question about the size of different bases. Since bases may be infinite, we have to be careful about what we mean. If $S$ and $T$ are sets, then we say that $S$ and $T$ have the same cardinality if there is a bijection between $S$ and $T$. If $S$ and $T$ are finite, this is equivalent to saying $S$ and $T$ have the same number of elements. While we won’t prove it here, it is true that if $B$ and $B'$ are bases for a vector space, then $B$ and $B'$ have the same cardinality.