1 Introductory Material

1.1 Properties & Definitions

Notation concerns: let V be a vector space with U a subspace of V: $U \leq V$ (the $\leq$ is for relation between spaces). Let C be a subset of V; this is written as $C \subset V$.
C is therefore a set of vectors, all in V. There are some things we can say about C (this is review material from basic linear algebra):

- C is *independent*: the sum $\sum_1^k \alpha_i c_i = \theta$ (the zero vector) $\implies \alpha_i = 0$ for all i, where c_i are all of the vectors in C. In prose, any finite linear combination equalling θ implies that the scalars in the linear combination are all 0; i.e. it is the trivial combination.
- C *spans* V: $\forall v \in V \; \exists \alpha_i \mid \sum_1^k \alpha_i c_i = v$. In prose, every vector in the space V is a linear combination of vectors in the span, C.
- C is a *basis* of V: C is a linearly independent spanning set of V.

Notice that these all assume C is a finite subset. This is an important distinction: we will discuss it after Zorn's Lemma but for now be aware that these are only definitions for finite subsets.

1.2 Theorems

Another property of basis sets is that they are maximally independent in V (independence is a quality of the subset, but to be maximally so depends on the space the subset is taken from).

Theorem 1. $C \subset V$ is a basis $\Leftrightarrow$ C is maximally independent in V.
C is maximally independent in V means that for every D with $V \supset D \supset C$, D is not linearly independent (in other words, D is linearly dependent, or that there is a nontrivial combination of vectors from D that add up to θ).

Proof. $\Leftarrow$: Let C be maximally independent in V. Let v be a vector not in C, call $D = \{v\} \cup C$. D is dependent by assumption; hence $\exists \sum_1^k \alpha_i d_i = \theta$ where not all $\alpha_i = 0$. One of the d_i vectors must be v since C is itself linearly independent (else we would have a nontrivial combination of vectors only from C that equal the zero vector, which violates our assumption); without loss of generality assume d_1 is v (we can reindex the finite sum any way we please). In other words, $\theta = \alpha_1 v + \sum_2^k \alpha_i d_i$; hence $v = -\sum_2^k \frac{\alpha_i d_i}{\alpha_1}$ for any v; d_i in the sum is now only vectors from C since the only vector not in C was v. This means that any vector $v \in V$ can be rewritten as a linear combination of vectors from C; C is a basis.

$\Rightarrow$: Let $v \neq \theta \in V$; hence C is a basis means there is a linear combination $\sum_{1}^{k} \alpha_i c_i = v$. Let $D \supset C$ with $v \in D$ and $v \not\in C$; thus the linear combination $-v + \sum_{1}^{k} \alpha_i c_i = \theta$ implies that D is linearly dependent; C is maximally independent. $\square$

So far there has been one result lurking in the background behind all of this work. We would be poor algebraists indeed if we did not set ourselves the task of the following theorem:

Theorem 2. Every vector space has a basis.

In order to prove this result we will employ a bit of artifice known as Zorn's Lemma.

1.3 Zorn's Lemma

This section will be a discussion of the statement, meaning, and use of Zorn's Lemma.

Zorn's Lemma: Let X be a partially ordered set such that every chain in X has a least upper bound; then X has maximal elements.

The lemma was originally proposed as a new axiom of set theory; it has many applications inside algebra. As it turns out, the lemma is equivalent to the axiom of choice (and the well-ordering theorem). As a result, there is no proof so much as a justification of its use as an axiom. The set-theoretic implications are outside the scope of the class; however we will need to understand the use of the lemma for linear algebra proofs.

Given a set $X \neq \emptyset$, X is *partially ordered* if there exists a binary relation $\leq$ on X such that:

- (Reflexivity) $x \leq x$
- (Antisymmetry) $x \leq y, y \leq x \implies x = y$
- (Transitivity) $x \leq y, y \leq z \implies x \leq z$

A *full* or *total* order is one where for any x and y in X we can say $x \leq y$ or $y \leq x$. An example of a partial order (as distinct from a total order) is the order given by the inclusion relation on the power set of X for nonempty X. Consider $1, 2, 3$ and its power set and how the different sets are included in each other to see the antisymmetric, transitive, and partial properties.

A *chain* Y in X is a totally ordered sequence of subsets $C_1 \subset C_2 \subset ...$ of X. Note that a chain can comprise of infinite sets; as an example, consider the power set of the integers and the inclusion relation. The set $\{I_n\}$ where n is a positive integer and I_n is the set of all integers less than n is a chain in the ordering. This interpretation helps understand the following terms defined on chains: we can interpret infinite sequences as chains.

The remaining terms of the hypothesis of Zorn's Lemma are:

- $M \in X$ is a maximum of Y if for any $N \supset M$ then $N \not\in Y$
- An upper bound of Y is an element $B \in X$ such that $B \geq C\ \forall C \in Y$ (keep track of the distinction between the inclusion relation for sets in the chain and the order relation $\geq$)
- The least upper bound is the smallest element L in the set of all bounds B for a given chain Y.

We are now in the position to prove the theorem from the previous section.

1.4 Linear Algebra

Theorem 2. Every vector space has a basis.

Proof. Don't blink or you'll miss it.
Let V be a vector space and X be a family of independent subsets of V. X is partially ordered under the set inclusion relation; let $C_1 \subset C_2 ... \subset C_n \subset ...$ be a chain in X; consider the union $C = \bigcup_i C_i$ (countable or finite). C is certainly a bound on the chain by the union property; we will show it is the least upper bound. Suppose D is an upper bound; thus $C_i \in D\ \forall i$; hence $\bigcup_i C_i \in D$, hence $C \in D$. Thus for any bound D for the chain, $C \leq D$. Thus every chain in X has a least upper bound; hence by Zorn's Lemma X has maximal elements; this is a basis for X. □

Note that this result does not tell us how to find the basis, or in fact anything useful at all other than "every vector space has a basis". For homework next Thursday (Feb. 6), adapt the technique of this proof to show that any independent subset of a vector space is contained in a basis of that vector space.

A side implication of basis sets and vector spaces is that any two basis of the same space must have the same cardinality (this result is trivial for finite basis sets). The theorem as proved applies not only to finite but to infinite vector spaces; the collection of all infinite sequences, for example. Intriguingly we know that there is a basis, but we have no idea at present what it looks like or anything about it. Let us bring in the concept of homomorphism before we discuss other examples of vector spaces.
If V is a vector space and W is another vector space, the map $L : V \to W$ is a linear transformation (or linear) if for $v \in V$ and $w \in W$

- $L(v + w) = L(v) + L(w)$
- $L(\alpha v) = \alpha L(v)$

Sometimes "L is linear" is written as $L \in \mathrm{Hom}(V, W)$.

Theorem 3. Given any vector $v \in V$, the basis representation of the vector v is unique.

Proof. Let B be the basis of V; suppose for sake of contradiction two different linear combinations that add up to v: $\sum \alpha_i b_i = v$ and $\sum \beta_i b_i = v$ where for at least one i, $\alpha_i \neq \beta_i$. Thus the difference $\sum (\alpha_i - \beta_i) b_i = \theta$ is a nontrivial linear combination to the zero vector, which implies the basis B is linearly dependent; a contradiction. □

As a result of this, since every vector can be uniquely written as a combination of the basis vectors, a linear transformation L is completely determined by the values it takes of the basis vectors. In other words, we can speak of vector spaces and homomorphisms between them entirely in terms of basis elements and their images. The function $f : B \to W$ from the basis of V to W where $f(v_i) = w_i$ gives rise to the linear transformation $L : V \to W$ where for any v $L(v) = L(\sum \alpha_i v_i) = \sum \alpha_i f(v_i)$.

1.4.1 Example: L^2 Space

In the L^2 space V here is the space of all polynomials on a set $U \subset \mathbb{R}^n$ (n-variable polynomials). Thus there is a linear transformation $L : V \to V$ defined by $L(p) = \nabla^2 p = \nabla \cdot \nabla p$ (the divergence of the gradient of the polynomial). We know this is a linear transformation because differentiation is a linear operator. Looking ahead, the kernel of L happens to be the set of harmonic functions. The basis, while it exists because of the previous theorem, is not easily known. Contemplate the meaning of Taylor series, Fourier series in the context of linear algebra.