Bases, etc. *(Sec. 9)*

NOTATION: $\vec{v}$, etc., $\in \mathcal{V}$, λ, etc., $\in \mathcal{F}$

Definition: A finite sum $\sum_{j=1}^{N} \lambda^j \vec{v}_j$ is a *linear combination* of the vectors $\{\vec{v}_j\}$. ($\lambda^j = 0$ is allowed.)

REMARK: In this course we have no concept of convergence, so we don't consider infinite sums — *except* in the sense that every finite sum can be thought of as an infinite sum with all but finitely many $\lambda^j = 0$. That convention will prove useful.

Definition: A set of vectors (possibly infinite) is *linearly dependent* if there is some linear combination of (distinct) vectors taken from the set, with at least one coefficient λ^j not equal to 0, which equals the zero vector. [A sequence $\{\vec{v}_j\}$ of vectors (possibly infinite) is called linearly dependent if there is some (finite) linear combination of the $\vec{v}_j$ with at least one coefficient λ^j not equal to 0, which equals the zero vector.]

This makes possible the following more important definition:

Definition: A set [or sequence] $\mathcal{S}$ is *linearly independent* if it's not dependent; that is, equations of the form

$$\sum \lambda^j \vec{v}_j = 0$$

involving vectors $\vec{v}_j$ from $\mathcal{S}$ are always false except in the trivial case that *all* the λ^j are 0.

Usually the set or sequence considered has 2 or more elements. In that case, *dependence* is equivalent to:

One of the vectors in the set is a linear combination of the others.

(Pick a nonzero λ^j and solve for $\vec{v}_j$.) In particular, any sequence with a repetition ($\vec{v}_j = \vec{v}_k$ for some $j \neq k$) is dependent.

To keep our logic sound, let's ask ourselves what happens when the number of elements in $\mathcal{S}$ is 0 or 1:

$\{\vec{v}\}$ is independent if $\vec{v} \neq 0$. [cf. Thm. 8.1]
$\{\vec{0}\}$ is dependent.
$\emptyset$ (the empty set) is independent. [different convention from Bowen & Wang]

REMARK: Recall that a *sequence* differs from a *set* in two ways: (1) the order of the elements matters; (2) repetitions are allowed. Every sequence determines a unique set, but the converse is false. Although Bowen & Wang (and most other books) frame the definition of linear independence in terms of sets, a case can be made that the corresponding notion for sequences is more important and natural. For instance, suppose that we are handed two vectors (defined, say, by two formulas, or by two rows of a matrix), and we want to determine whether they are linearly independent. If the two vectors turn out to be *the same*, we surely want the answer to be "no"! (If this is not clear now, it should become clear in the light of the next theorem.) But it is the two-element sequence $\{\vec{v}, \vec{v}\}$ which is dependent, not the one-element set $\{\vec{v}\}$ it determines. Strictly speaking, every definition and every theorem comes in two versions, one for sets and one for sequences. But I shall usually state only one of them.

Theorem 9.7′. *If the sequence $\{\vec{v}_j\}_{j=1}^N$ is linearly independent, then the coefficients in any (true) equation of the form*

$$\vec{u} = \sum_{j=1}^{N} \lambda^j \vec{v}_j$$

are unique.

PROOF: Suppose that also $\vec{u} = \sum_{j=1}^N \mu^j \vec{v}_j$. Then $\vec{0} = \sum_{j=1}^N (\lambda^j - \mu^j)\vec{v}_j$, hence $\lambda^j - \mu^j = 0$ for all j, so $\mu^j = \lambda^j$ after all.

REMARK: The theorem remains valid for an *infinite* linearly independent sequence, provided we interpret a finite linear combination as an infinite sum in which all but finitely many λ^j are 0. Note that if, instead, one formulates such things in terms of finite subsequences (as Milne does, for instance), one has to worry about the possibility of sums which differ only by terms with coefficients $\lambda = 0$. This would make Thm. 9.7, for instance, technically untrue.

PROCEDURAL REMARK: Theorem references are to Bowen and Wang; often I will have changed the statement of the theorem considerably. If the change involves a significant change in the *content* of the proposition, I put a prime ("′") on the theorem number. Sometimes the unprimed theorem will also show up (as in the present case — see below). I shall try to label the more important theorems by typing "Theorem" in **boldface**, whereas for lemmas encountered in passing, the word will be in SMALL CAPS.

THEOREM (9.1, 9.3). *If $\mathcal{S}$ is dependent, then any superset of $\mathcal{S}$ is dependent.* [Why?] *Hence if $\mathcal{S}$ is independent, any subset of it is independent.*

THEOREM 9.2. *Any set containing $\vec{0}$ is dependent.* [Why?]

Definition: A set $\mathcal{S}$ *spans* (or *generates*) $\mathcal{V}$ if every $\vec{u} \in \mathcal{V}$ can be written as a (finite) linear combination, $\vec{u} = \sum_{j=1}^N \lambda^j \vec{v}_j$, with $\vec{v}_j \in \mathcal{S}$ for all j.

Definition: For any set $\mathcal{S} \subset \mathcal{V}$, the *span* of $\mathcal{S}$ (written $\operatorname{span} \mathcal{S}$) is the set of all linear combinations of members of $\mathcal{S}$. (If $\mathcal{S} = \emptyset$, then $\operatorname{span} \mathcal{S} \equiv \{\vec{0}\}$.)

Remark: $\operatorname{span}\mathcal{S}$ is always a subspace (since it's algebraically closed). If $\mathcal{S}$ is already a subspace, then $\operatorname{span}\mathcal{S} = \mathcal{S}$.

THEOREM 9.9. *If $\mathcal{S}$ spans $\mathcal{V}$, then any superset of $\mathcal{S}$ spans $\mathcal{V}$.* [Why?]

Finally, the big definition that all this has been leading up to:

Definition: $\{\vec{v}_j\} \subset \mathcal{V}$ is a *basis* for $\mathcal{V}$ if it is linearly independent [as a sequence!] *and* it spans $\mathcal{V}$.

Theorem 9.7. *If $\{\vec{v}_j\}$ is a basis, then every $\vec{u} \in \mathcal{V}$ can be expressed as*

$$\vec{u} = \sum_{j=1}^{N} \lambda^j \vec{v}_j$$

in **exactly one way**. [See remark below Thm. 9.7′.]

PROOF: Existence $\Leftarrow$ definition of "span". Uniqueness $\Leftarrow$ Theorem 9.7′.

Examples of bases:

1. The usual unit vectors in 3-dimensional space, $\{\hat{\imath}, \hat{\jmath}, \hat{k}\}$.

2. The power functions $\{t^n\}$ form a basis for the vector space of polynomials.

3. A basis for the space of (complex) 2×2 matrices is

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

4. Another, very useful, basis for the same space is

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}.$$

   These (especially the last three) are called the *Pauli matrices* (by physicists, at least); they play a key role in the study of the group of rotations in 3-dimensional space.

5. $\mathcal{V}$ = solution space of $\frac{d^2y}{dx^2} + \omega^2 y = 0 \quad (y \in \mathbf{C},\ x \in \mathbf{R})$. One basis is $\{e^{i\omega x}, e^{-i\omega x}\}$. Another is $\{\cos\omega x, \sin\omega x\}$.

Theorem 9.4. *Any two bases for the same vector space contain the same number (possibly ∞) of vectors.* [proof postponed]

Definition: The number of vectors in a basis of $\mathcal{V}$ is the *dimension* of $\mathcal{V}$ (abbreviated $\dim \mathcal{V}$).

REMARK: In this course we shall not distinguish among different kinds of "infinity" (i.e., transfinite cardinal numbers). All our statements about bases and dimension involving infinite-dimensional spaces will be *correct*, but they will be *incomplete* because of our failure to introduce the finer classification of infinite-dimensional spaces based on the cardinality of their bases.

EXAMPLES:

$\mathbf{R}^n$ and $\mathbf{C}^n$ have dimension n.

$\mathbf{C}^n$ as a *real* vector space has dimension $2n$.

The solution space of our ODE (example 5 above) has dimension 2.

The solution space of $\nabla^2\phi = 0$ has infinite dimension. (This space is not well defined until I state a domain for the functions ϕ. Let's say $\phi = \phi(x, y)$ where $(x, y) \in$ some rectangle $B \subset \mathbf{R}^2$.)

$\{\vec{0}\}$ is a 0-dimensional vector space with basis $\emptyset$.

A 1-dimensional space consists of all scalar multiples of some vector $\vec{v} \neq 0$, and $\{\vec{v}\}$ is a basis for it.

Theorem 9.4 is a corollary of:

EXCHANGE LEMMA. [statement and proof adapted from Milne, pp. 30–31; Bowen & Wang Thm. 9.8 is slightly different.] *If $S \subset \mathcal{V}$ spans $\mathcal{V}$, and $T \subset \mathcal{V}$ is linearly independent, then s, the number of vectors in S, is greater than or equal to t, the number of vectors in T. Furthermore, a new spanning set can be constructed in which t elements of S have been replaced by the elements of T [at least if t is finite].*

PROOF OF THEOREM: Let B and B' be bases containing n and n' vectors, respectively.

$$S = B \text{ and } T = B' \Rightarrow n \geq n'.$$

$$S = B' \text{ and } T = B \Rightarrow n' \geq n.$$

Therefore $n = n'$.

Proof of Lemma: If s and t are both ∞, clause 1 is true and clause 2 is vacuous. For expository clarity assume S and T countable. Let $T \equiv \{\vec{y}_1, \vec{y}_2, \ldots, \vec{y}_j, \ldots\}$, and $S \equiv \{\vec{x}_1, \vec{x}_2, \ldots\}$. Since S spans, $\vec{y}_1$ is a linear combination of the $\vec{x}$'s. Since T is independent, $\vec{y}_1 \neq \vec{0}$; hence in the expansion $\vec{y}_1 = \sum \lambda^j \vec{x}_j$ at least one λ^j is nonzero. Solve for the first such $\vec{x}_j$ in terms of $\vec{y}_1$ and the other $\vec{x}$'s. Get a new spanning set S' by replacing $\vec{x}_j$ in S by $\vec{y}_1$. [Why does this new set still span?] Now $\vec{y}_2$ is a linear combination of vectors in S'. Since T is independent, $\vec{y}_2$ is not a multiple of $\vec{y}_1$; hence $\vec{y}_2$ contains at least one term proportional to an $\vec{x}$. Solve for the first such $\vec{x}$, and replace it by $\vec{y}_2$ to get a new spanning set, S''. Continue in this way until you run out of vectors in either S or T.

Case I: $s < t$. We have exhausted S but have vectors in T left over (say $\vec{y}_{491}, \ldots$). Then $\vec{y}_{491} \in \operatorname{span} S''''\cdots$ (490 primes), where $S''''\cdots = \{\vec{y}_1, \ldots, \vec{y}_{490}\}$ (not necessarily in this order). This contradicts linear independence of T! So this case doesn't occur (which is the main assertion of the lemma).

Case II: $s \geq t$. Clause 1 is true, and we have succeeded in exchanging all of T for elements of S, which is clause 2.

Corollary. *Let $\mathcal{V}$ be n-dimensional, n* **finite**. *Let $\mathcal{S} \equiv \{\vec{v}_j\}_{j=1}^{M} \subset \mathcal{V}$. Then:*

(1) $M < n \Rightarrow \mathcal{S}$ does not span. (It may or may not be independent.)

(2) $M > n \Rightarrow \mathcal{S}$ is not linearly independent. (It may or may not span.)

(3) $M = n \Rightarrow \mathcal{S}$ is independent if and only if it spans. (So to show that $\mathcal{S}$ is a **basis**, *it suffices to check just* **one** *of the two clauses in the definition of basis.)*

Proof: Compare $\mathcal{S}$ with a basis $\mathcal{B}$, using the exchange lemma. [Cf. proof of Thm. 9.8′ below.]

Visualization of (3): In $\mathbf{R}^3$, three vectors either are coplanar or they're not. Contemplate the meaning of "span" and "independent" in this situation.

Theorem 9.8′. *If $\mathcal{V}$ is finite-dimensional,*

(1) Any [finite] spanning set can be made into a basis by discarding vectors (if necessary).

(2) Any linearly independent set can be made into a basis by adding vectors (if necessary).

Proof:

(1) [Kolman, p. 66 (1970 ed.)] Throwing out a vector which is linearly dependent on the others does not change the span. Throw out such vectors one by one until only $\dim \mathcal{V}$ are left.

(2) Apply the exchange lemma to embed the independent set (in role of T) into a basis (in role of S). The resulting set spans (by the exchange lemma) and since it has $\dim \mathcal{V}$ elements it is a basis.

REMARK: In the infinite-dimensional case, recall that by our definitions, each vector can be expressed as a *finite* sum over the basis vectors. Such a basis is called a *Hamel basis* (or *algebraic basis*). Same remark for "span": The span of $\{t^n\}$ is the polynomials, not the analytic functions; the span of $\{e^{inx}\}$ is the "trigonometric polynomials", not the most general convergent Fourier series. The second example shows that in analysis we are likely to be interested in another kind of basis, such that each vector (in some function space) equals an *infinite* linear combination of basis elements, converging in some topology. (For present purposes, "topology" $\equiv$ "definition of convergence".) Such a basis is sometimes called a *Schauder basis*. The relative uselessness of infinite Hamel bases is why we haven't studied them in depth; for present purposes, the only thing we can say about an infinite-dimensional space is that it isn't finite-dimensional. In particular, we have not proved the following (true) theorem:

Every vector space has a (Hamel) basis (possibly uncountable). In other words, if $\mathcal{V}$ is not finite-dimensional, then it is infinite-dimensional.