Chapter 3

Vector Spaces

Immer mit den einfachsten Beispielen anfangen.

David Hilbert

1. REAL VECTOR SPACES

The basic models for vector spaces are the spaces of n-dimensional row or column vectors:

$\mathbb{R}^n$: the set of row vectors $v = (a_1, \ldots, a_n)$, or

the set of column vectors $v = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}$.

Though row vectors take less space to write, the definition of matrix multiplication makes column vectors more convenient for us. So we will work with column vectors most of the time. To save space, we will occasionally write a column vector in the form $(a_1, \ldots, a_n)^t$.

For the present we will study only two operations:

(1.1) *vector addition:*
$$\begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} + \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} a_1+b_1 \\ \vdots \\ a_n+b_n \end{bmatrix}, \text{ and}$$

scalar multiplication:
$$c\begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} ca_1 \\ \vdots \\ ca_n \end{bmatrix}.$$

These operations make $\mathbb{R}^n$ into a *vector space*. Before going to the formal definition of a vector space, let us look at some other examples—nonempty subsets of $\mathbb{R}^n$ closed under the operations (1.1). Such a subset is called a *subspace*.

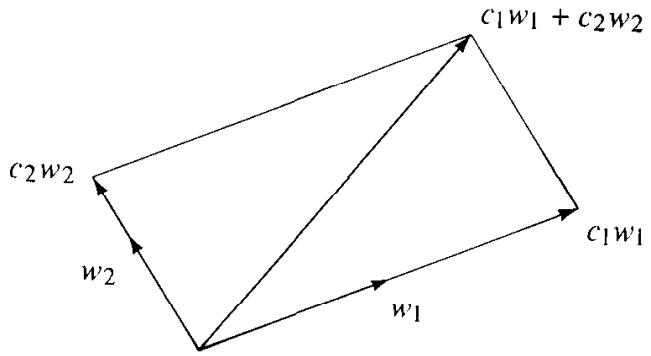
(1.2) **Example.** The subspaces W of the space $\mathbb{R}^2$ are of three types:

(i) the zero vector alone: $W = \{0\}$;
(ii) the vectors lying on a line L through the origin;
(iii) the whole space: $W = \mathbb{R}^2$.

This can be seen from the parallelogram law for addition of vectors. If W contains two vectors w_1, w_2 not lying on one line, then every vector v can be obtained from these two vectors as a "linear combination"

$$c_1w_1 + c_2w_2,$$

where c_1, c_2 are scalars. So $W = \mathbb{R}^2$ in this case. If W does not contain two such vectors, then we are in one of the remaining cases. □



Similarly, it can be shown that the subspaces of $\mathbb{R}^3$ are of four types:

(i) the zero vector;
(ii) the vectors lying on a line through the origin;
(iii) the vectors lying in a plane through the origin;
(iv) the whole space $\mathbb{R}^3$.

This classification of subspaces of $\mathbb{R}^2$ and $\mathbb{R}^3$ will be clarified in Section 4 by the concept of *dimension*.

Systems of homogeneous linear equations furnish many examples. The set of solutions of such a system is always a subspace. For, if we write the system in matrix notation as $AX = 0$, where A is an $m \times n$ matrix and X is a column vector, then it is clear that

(a) $AX = 0$ and $AY = 0$ imply $A(X + Y) = 0$. In other words, if X and Y are solutions, so is $X + Y$.
(b) $AX = 0$ implies $AcX = 0$: If X is a solution, so is cX.

For example, let W be the set of solutions of the equation

$$2x_1 - x_2 - 2x_3 = 0, \text{ or } AX = 0, \tag{1.3}$$

where $A = [2 \ -1 \ \ 2]$. This space is the set of vectors lying in the plane through the origin and orthogonal to A. Every solution is a linear combination $c_1w_1 + c_2w_2$ of two particular solutions w_1, w_2. Most pairs of solutions, for example

$$w_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \ w_2 = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \tag{1.4}$$

will span the space of solutions in this way. Thus every solution has the form

$$c_1w_1 + c_2w_2 = \begin{bmatrix} c_1+c_2 \\ 2c_2 \\ c_1 \end{bmatrix}, \tag{1.5}$$

where c_1, c_2 are arbitrary constants. Another choice of the particular solutions w_1, w_2 would result in a different but equivalent description of the space of all solutions.

(1.6) **Definition.** A *real vector space* is a set V together with two laws of composition:

(a) *Addition:* $V \times V \longrightarrow V$, written $v, w \rightsquigarrow v + w$

(b) *Scalar multiplication:* $\mathbb{R} \times V \longrightarrow V$, written $c, v \rightsquigarrow cv$

These laws of composition must satisfy the following axioms:

(i) Addition makes V into an abelian group V^+.

(ii) Scalar multiplication is associative with multiplication of real numbers:

$$(ab)v = a(bv).$$

(iii) Scalar multiplication by the real number 1 is the identity operation:

$$1v = v.$$

(iv) Two distributive laws hold:

$$(a + b)v = av + bv$$

$$a(v + w) = av + aw.$$

Of course all the axioms should be quantified universally; that is, they are assumed to hold for all $a, b \in \mathbb{R}$ and all $v, w \in V$.

The identity element for the addition law in V is denoted by 0, or by 0_V if there is danger of confusing the zero vector with the number zero.

Notice that scalar multiplication associates to every pair consisting of a real number c and a vector v another vector cv. Such a rule is called an *external law of composition* on the vector space.

Multiplication of two vectors is not a part of the structure, though various products, such as the cross product of vectors in $\mathbb{R}^3$, can be defined. These products aren't completely intrinsic; they depend on choosing coordinates. So they are considered to be additional structure on the vector space.

Read axiom (ii) carefully. The left side means multiply a and b as real numbers, then scalar multiply ab and v, to get a vector. On the right side, both operations are scalar multiplication.

The two laws of composition are related by the essential distributive laws. Note that in the first distributive law the symbol $+$ on the left stands for addition of real numbers, while on the right, it stands for addition of vectors.

(1.7) **Proposition.** The following identities hold in a vector space V:

(a) $0_{\mathbb{R}}v = 0_V$, for all $v \in V$,
(b) $c0_V = 0_V$, for all $c \in \mathbb{R}$,
(c) $(-1)v = -v$, for all $v \in V$.

Proof. To see (a), we use the distributive law to write

$$0v + 0v = (0 + 0)v = 0v = 0v + 0.$$

Cancelling $0v$ from both sides, we obtain $0v = 0$. Please go through this carefully, noting which symbols 0 refer to the number and which refer to the vector.

Similarly, $c0 + c0 = c(0 + 0) = c0$. Hence $c0 = 0$. Finally,

$$v + -1v = 1v + -1v = (1 + -1)v = 0v = 0.$$

Hence $-1v$ is the additive inverse of v. □

(1.8) **Examples.**

(a) A subspace of $\mathbb{R}^n$ is a vector space, with the laws of composition induced from those on $\mathbb{R}^n$.
(b) Let $V = \mathbb{C}$ be the set of complex numbers. Forget multiplication of complex numbers, and keep only addition $\alpha + \beta$ and multiplication $c\alpha$ of a complex number α by a real number c. These operations make $\mathbb{C}$ into a real vector space.
(c) The set of real polynomials $p(x) = a_n x^n + \cdots + a_0$ is a vector space, with addition of polynomials and multiplication of polynomials by scalars as its laws of composition.
(d) Let V be the set of continuous real-valued functions on the interval $[0, 1]$. Look only at the operations of addition of functions $f + g$ and multiplication of functions by numbers cf. This makes V a real vector space.

Note that each of our examples has more structure than we look at when we view it as a vector space. This is typical. Any particular example is sure to have some extra features which distinguish it from others, but this is not a drawback of the definition. On the contrary, the strength of the abstract approach lies in the fact that consequences of the general axioms can be applied to many different examples.

2. ABSTRACT FIELDS

It is convenient to treat the real and complex cases simultaneously in linear algebra. This can be done by listing the properties of the "scalars" which are needed axiomatically, and doing so leads to the notion of a *field*.

It used to be customary to speak only of subfields of the complex numbers. A *subfield* of $\mathbb{C}$ is any subset which is closed under the four operations addition, subtraction, multiplication, and division, and which contains 1. In other words, F is a subfield of $\mathbb{C}$ if the following properties hold:

(2.1)

(a) If $a, b \in F$, then $a + b \in F$.
(b) If $a \in F$, then $-a \in F$.
(c) If $a, b \in F$, then $ab \in F$.
(d) If $a \in F$ and $a \neq 0$, then $a^{-1} \in F$.
(e) $1 \in F$.

Note that we can use axioms (a), (b), and (e) to conclude that $1 - 1 = 0$ is an element of F. Thus F is a subset which is a subgroup of $\mathbb{C}^+$ under addition and such that $F - \{0\} = F^\times$ is a subgroup of $\mathbb{C}^\times$ under multiplication. Conversely, any such subset is a subfield.

Here are some examples of subfields of $\mathbb{C}$:

(2.2) **Examples.**

(a) $F = \mathbb{R}$, the field of real numbers.
(b) $F = \mathbb{Q}$, the field of rational numbers (= fractions of integers).
(c) $F = \mathbb{Q}[\sqrt{2}]$, the field of all complex numbers of the form $a + b\sqrt{2}$, where $a, b \in \mathbb{Q}$.

It is a good exercise to check axioms (2.1) for the last example.

These days, it is customary to introduce fields abstractly. The notion of an abstract field is harder to grasp than that of a subfield of $\mathbb{C}$, but it contains important new classes of fields, including finite fields.

(2.3) **Definition.** A *field* F is a set together with two laws of composition

$$F \times F \xrightarrow{+} F \quad \text{and} \quad F \times F \xrightarrow{\times} F$$

$$a, b \rightsquigarrow a + b \qquad a, b \rightsquigarrow ab$$

called addition and multiplication, and satisfying the following axioms:

(i) Addition makes F into an abelian group F^+. Its identity element is denoted by 0.

(ii) Multiplication is associative and commutative and makes $F^\times = F - \{0\}$ into a group. Its identity element is denoted by 1.

(iii) Distributive law: For all $a, b, c \in F$, $(a + b)c = ac + bc$.

The first two axioms describe properties of the two laws of composition, addition and multiplication, separately. The third axiom, the distributive law, is the one which relates addition to multiplication. This axiom is crucial, because if the two laws were unrelated, we could just as well study each of them separately. Of course we know that the real numbers satisfy these axioms, but the fact that they are all that is needed for arithmetic operations can only be understood after some experience in working with them.

One can operate with matrices A whose entries a_{ij} are in any field F. The discussion of Chapter 1 can be repeated without change, and you should go back to look at this material again with this in mind.

The simplest examples of fields besides the subfields of the complex numbers are certain finite fields called the prime fields, which we will now describe. We saw in Section 9 of Chapter 2 that the set $\mathbb{Z}/n\mathbb{Z}$ of congruence classes modulo n has laws of addition and multiplication derived from addition and multiplication of integers. Now all of the axioms for a field hold for the integers, except for the existence of multiplicative inverses in axiom (2.3ii). The integers are not closed under division. And as we have already remarked, such axioms carry over to addition and multiplication of congruence classes. But there is no reason to suppose that multiplicative inverses will exist for congruence classes, and in fact they need not. The class of 2, for example, does not have a multiplicative inverse modulo 6. So it is a surprising fact that if p is a prime integer then all nonzero congruence classes modulo p have inverses, and therefore the set $\mathbb{Z}/p\mathbb{Z}$ is a field. This field is called a *prime field* and is usually denoted by $\mathbb{F}_p$:

$$\mathbb{F}_p = \{\bar{0}, \bar{1}, \ldots, \overline{p-1}\} = \mathbb{Z}/p\mathbb{Z}. \tag{2.4}$$

(2.5) **Theorem.** Let p be a prime integer. Every nonzero congruence class $\bar{a}$ (modulo p) has a multiplicative inverse, and hence $\mathbb{F}_p$ is a field with p elements.

The theorem can also be stated as follows:

(2.6) *Let p be a prime, and let a be any integer not divisible by p. There is an integer b such that $ab \equiv 1$* (modulo p).

For $ab \equiv 1$ (modulo p) is the same as $\overline{a}\overline{b} = \overline{ab} = \overline{1}$, which means that $\overline{b}$ is the multiplicative inverse of $\overline{a}$.

For example, let $p = 13$ and $\overline{a} = \overline{6}$. Then $\overline{a}^{-1} = \overline{11}$ because

$$6 \cdot 11 = 66 \equiv 1 \text{ (modulo 13)}.$$

Finding the inverse of a congruence class $\overline{a}$ (modulo p) is not easy in general, but it can be done by trial and error if p is small. A systematic way is to compute the powers of $\overline{a}$. Since every nonzero congruence class has an inverse, the set of all of them forms a finite group of order $p - 1$, usually denoted by $\mathbb{F}_p^{\times}$. So every element $\overline{a}$ has finite order dividing $p - 1$. Thus if $p = 13$ and $\overline{a} = \overline{3}$, we find $\overline{a}^2 = \overline{9}$, and $\overline{a}^3 = \overline{27} = \overline{1}$, which shows that $\overline{a}$ has order 3. We are lucky: $\overline{a}^{-1} = \overline{a}^2 = \overline{9}$. On the other hand, if we had tried this method with $\overline{a} = \overline{6}$, we would have found that $\overline{6}$ has order 12. The computation would have been lengthy.

Proof of Theorem (2.5). Let $\overline{a} \in \mathbb{F}_p$ be any nonzero element, and let us use the method just discussed to show that $\overline{a}$ has an inverse. We consider the powers $1, \overline{a}, \overline{a}^2, \overline{a}^3, \ldots$. Since there are infinitely many powers and only finitely many elements in $\mathbb{F}_p$, there must be two powers which are equal, say $\overline{a}^m = \overline{a}^n$, where $m < n$. At this point, we would like to cancel $\overline{a}^m$, to obtain $\overline{1} = \overline{a}^{n-m}$. Once this cancellation is justified, we will have shown that $\overline{a}^{n-m-1}$ is the inverse of $\overline{a}$. This will complete the proof.

Here is the cancellation law we need:

(2.7) **Lemma.** *Cancellation Law:* Let $\overline{a}, \overline{c}, \overline{d}$ be elements of $\mathbb{F}_p$ with $\overline{a} \neq \overline{0}$. If $\overline{a}\overline{c} = \overline{a}\overline{d}$, then $\overline{c} = \overline{d}$.

Proof. Set $\overline{b} = \overline{c} - \overline{d}$. Then the statement of the lemma becomes: If $\overline{a}\overline{b} = \overline{0}$ and $\overline{a} \neq 0$, then $\overline{b} = 0$. To prove this, we represent the congruence classes $\overline{a}, \overline{b}$ by integers a, b. Then what has to be shown is the following intuitively plausible fact:

(2.8) **Lemma.** Let p be a prime integer and let a, b be integers. If p divides the product ab, then p divides a or p divides b.

Proof. Suppose that p does not divide a, but that p divides ab. We must show that p divides b. Since p is a prime, 1 and p are the only positive integers which divide it. Since p does not divide a, the only common divisor of p and a is 1. So 1 is their greatest common divisor. By Proposition (2.6) of Chapter 2, there are integers r, s so that $1 = rp + sa$. Multiply both sides by b: $b = rpb + sab$. Both of the terms on the right side of this equality are divisible by p; hence the left side a is divisible by p too, as was to be shown. □

As with congruences in general, computations in the field $\mathbb{F}_p$ can be made by working with integers, except that division can not be carried out in the integers. This difficulty can often be handled by putting everything on a common denominator in such a way that the required division is left until the end. For example, suppose we ask for solutions of a system of n linear equations in n unknowns, in the field $\mathbb{F}_p$.

We represent the system of equations by an integer system, choosing representatives for the residue classes in a convenient way. Say that the integer system is $AX = B$, where A is an $n \times n$ integer matrix and B is an integer column vector. Then to solve the system in $\mathbb{F}_p$, we try to invert the matrix A modulo p. Cramer's Rule, (adj A)$A = \delta I$, where $\delta = \det A$, is a formula valid in the integers [Chapter 1 (5.7)], and therefore it also holds in $\mathbb{F}_p$ when the matrix entries are replaced by their congruence classes. If the residue class of δ is not zero, then we can invert the matrix A in $\mathbb{F}_p$ by computing δ^{-1}(adj A).

(2.9) **Corollary.** Consider a system $AX = B$ of n linear equations in n unknowns where the entries of A, B are in $\mathbb{F}_p$. The system has a unique solution in $\mathbb{F}_p$ if $\det A \neq 0$ in $\mathbb{F}_p$. □

For example, consider the system of linear equations $AX = B$, where

$$A = \begin{bmatrix} 8 & 3 \\ 2 & 6 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 3 \\ -1 \end{bmatrix}.$$

Since the coefficients are integers, they define a system of equations in $\mathbb{F}_p$ for any prime p. The determinant of A is 42, so the system has a unique solution in $\mathbb{F}_p$ for all p different from 2, 3 and 7. Thus if $p = 13$, we find $\det A = 3$ when evaluated (modulo 13). We already saw that $3^{-1} = 9$ in $\mathbb{F}_{13}$. So we can use Cramer's Rule to compute

$$A^{-1} = \begin{bmatrix} 2 & -1 \\ 8 & 7 \end{bmatrix} \quad \text{and} \quad X = A^{-1}B = \begin{bmatrix} 7 \\ 4 \end{bmatrix}, \text{ in } \mathbb{F}_{13}.$$

The system has no solution in $\mathbb{F}_2$ or $\mathbb{F}_3$. It happens to have solutions in $\mathbb{F}_7$, though $\det A = 0$ in that field.

We remark in passing that invertible matrices with entries in the field $\mathbb{F}_p$ provide new examples of finite groups—the general linear groups over finite fields:

$$GL_n(\mathbb{F}_p) = \{n \times n \text{ invertible matrices with entries in } \mathbb{F}_p\}.$$

The smallest of these is the group $GL_2(\mathbb{F}_2)$ of invertible 2×2 matrices with entries (modulo 2), which consists of the six matrices

(2.10)

$$GL_2(\mathbb{F}_2) = \left\{ \begin{bmatrix} 1 & \\ & 1 \end{bmatrix}, \begin{bmatrix} & 1 \\ 1 & \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & \end{bmatrix}, \begin{bmatrix} 1 & \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} & 1 \\ 1 & 1 \end{bmatrix} \right\}.$$

There is one property of the finite fields $F = \mathbb{F}_p$ which distinguishes them from subfields of $\mathbb{C}$ and which affects computations occasionally. This property is that adding 1 to itself a certain number of times (in fact p times) gives 0. A field F is said to have *characteristic* p if $1 + \cdots + 1$ (p terms) $= 0$ in F, and if p is the smallest positive integer with that property. In other words, the characteristic of F is the order of 1, as an element of the additive group F^+, provided that the order is finite (Chapter 2, Section 2). In case the order is infinite, that is, $1 + \cdots + 1$ is

never 0 in F, the field is, paradoxically, said to have *characteristic zero*. Thus subfields of $\mathbb{C}$ have characteristic zero, while the prime field $\mathbb{F}_p$ has characteristic p. It can be shown that the characteristic of any field F is either zero or a prime number.

Now let F be an arbitrary field. A vector space over a field F is defined as in (1.6), with F replacing $\mathbb{R}$.

(2.11) **Definition.** A *vector space* V over a field F is a set together with two laws of composition:

(a) *addition:* $V \times V \longrightarrow V$, written $v, w \rightsquigarrow v + w$,

(b) *scalar multiplication:* $F \times V \longrightarrow V$, written $c, v \rightsquigarrow cv$,

and satisfying the following axioms:

(i) Addition makes V into a commutative group V^+.

(ii) Scalar multiplication is associative with multiplication in F:

$$(ab)v = a(bv), \text{ for all } a, b \in F \text{ and } v \in V.$$

(iii) The element 1 acts as identity: $1v = v$, for all $v \in V$.

(iv) Two distributive laws hold:

$$(a + b)v = av + bv \quad \text{and} \quad a(v + w) = av + aw,$$

for all $a, b \in F$ and $v, w \in V$.

All of Section 1 can be repeated, replacing the field $\mathbb{R}$ by F. Thus the space F^n of row vectors $(a_1, \ldots, a_n)$, $a_i \in F$, is a vector space over F and so on.

It is important to note that the definition of vector space includes implicitly the choice of a field F. The elements of this field F are often called *scalars*. We usually keep this field fixed. Of course, if V is a complex vector space, meaning a vector space over the field $\mathbb{C}$, and if $F \subset \mathbb{C}$ is any subfield, then V is also naturally a vector space over F because cv is defined for all $c \in F$. But we consider the vector space structure to have changed when we restrict the scalars from $\mathbb{C}$ to F.

Two important concepts analogous to subgroups and isomorphisms of groups are the concepts of subspace and of isomorphism of vector spaces. We have already defined subspaces for complex vector spaces, and the definition is the same for any field. A *subspace* W of a vector space V (over a field F) is a subset with the following properties:

(2.12)

(a) If $w, w' \in W$, then $w + w' \in W$.

(b) If $w \in W$ and $c \in F$, then $cw \in W$.

(c) $0 \in W$.

A subspace W is called a *proper* subspace of V if it is neither the whole space V nor the zero subspace $\{0\}$.

It is easy to see that a subspace is just a subset on which the laws of composition induce the structure of vector space.

As in Section 1, the space of all solutions of a system of m linear equations in n unknowns

$$AX = 0,$$

with coefficients in F, is an example of a subspace of the space F^n.

(2.13) **Definition.** An *isomorphism* φ from a vector space V to a vector space V', both over the same field F, is a bijective map $\varphi: V \longrightarrow V'$ compatible with the laws of composition, that is, a bijective map satisfying

$$\text{(a) } \varphi(v + v') = \varphi(v) + \varphi(v') \quad \text{and} \quad \text{(b) } \varphi(cv) = c\varphi(v),$$

for all $v, v' \in V$ and all $c \in F$.

(2.14) **Examples.**

(a) The space F^n of n-dimensional row vectors is isomorphic to the space of n-dimensional column vectors.

(b) View the set of complex numbers $\mathbb{C}$ as a real vector space, as in (1.8b). Then the map $\varphi: \mathbb{R}^2 \longrightarrow \mathbb{C}$ sending $(a, b) \rightsquigarrow a + bi$ is an isomorphism.

3. BASES AND DIMENSION

In this section we discuss the terminology used when working with the two operations, addition and scalar multiplication, in an abstractly given vector space. The new concepts are *span, linear independence,* and *basis.*

It will be convenient to work with *ordered* sets of vectors here. The ordering will be unimportant much of the time, but it will enter in an essential way when we make explicit computations. We've been putting curly brackets around unordered sets, so in order to distinguish ordered from unordered sets, let us enclose ordered sets with round brackets. Thus the ordered set (a, b) is considered different from the ordered set (b, a), whereas the unordered sets $\{a, b\}$ and $\{b, a\}$ are considered equal. Repetitions will also be allowed in an ordered set. So (a, a, b) is considered an ordered set, and it is different from (a, b), in contrast to the convention for unordered sets, where $\{a, a, b\}$ would denote the same set as $\{a, b\}$.

Let V be a vector space over a field F, and let $(v_1, \ldots, v_n)$ be an ordered set of elements of V. A *linear combination* of $(v_1, \ldots, v_n)$ is any vector of the form

$$w = c_1v_1 + c_2v_2 + \cdots + c_nv_n, \; c_i \in F. \tag{3.1}$$

For example, suppose that the ordered set consists of the two vectors in $\mathbb{R}^3$ considered in (1.4): $v_1 = (1, 0, 1)^t$ and $v_2 = (1, 2, 0)^t$. Then a linear combination will have the form (1.5): $(c_1 + c_2, 2c_2, c_1)^t$. The vector $(3, 4, 1)^t = v_1 + 2v_2$ is one such linear combination.

A solution X of a system of linear equations written in the matrix form $AX = B$ [Chapter 1 (1.9)] exhibits the column vector B as a linear combination of the columns of the matrix A. The coefficients are the entries of the vector X.

A linear combination of a single vector (v) is just a multiple cv or v .

The set of all vectors w which are linear combinations of $(v_1, \ldots, v_n)$ forms a subspace W of V, called the subspace *spanned* by the set: If w (3.1) and $w' = c_1'v_1 + \cdots + c_n'v_n$ are elements of W, then so is

$$w + w' = (c_1 + c_1')v_1 + \cdots + (c_n + c_n')v_n,$$

and if $a \in F$, then $aw = (ac_1)v_1 + \cdots + (ac_n)v_n$ is in W. So $w + w'$ and aw are in W. Finally, $0 = 0v_1 + \cdots + 0v_n \in W$. This shows that the conditions of (2.12) hold.

The space spanned by a set S will often be denoted by Span S. Clearly, Span S is the smallest subspace of V which contains S. We could also call it the subspace *generated* by S. Note that the order is irrelevant here. The span of S is the same as the span of any reordering of S.

One can also define the span of an infinite set of vectors. We will discuss this in Section 5. In this section, let us assume that our sets are *finite*.

(3.2) **Proposition.** Let S be a set of vectors of V, and let W be a subspace of V. If $S \subset W$, then Span $S \subset W$.

This is obvious, because W is closed under addition and scalar multiplication. If $S \subset W$, then any linear combination of vectors of S is in W too. □

A *linear relation* among vectors $v_1, \ldots, v_n$ is any relation of the form

$$c_1v_1 + c_2v_2 + \cdots + c_nv_n = 0, \tag{3.3}$$

where the coefficients c_i are in F. An ordered set $(v_1, \ldots, v_n)$ of vectors is called *linearly independent* if there is no linear relation among the vectors in the set, except for the trivial one in which all the coefficients c_i are zero. It is useful to state this condition positively:

(3.4) *Let $(v_1, \ldots, v_n)$ be a linearly independent set. Then from the equation $c_1v_1 + \cdots + c_nv_n = 0$, we can conclude that $c_i = 0$ for every $i = 1, \ldots, n$.*

Conversely, if (3.4) holds, then the vectors are linearly independent.

The vectors (1.4) are linearly independent.

Note that a linearly independent set S can not have any repetitions. For if two vectors v_i, v_j of S are equal, then

$$v_i - v_j = 0$$

is a linear relation of the form (3.3), the other coefficients being zero. Also, no vector v_i of a linearly independent family may be zero, because if it is, then $v_i = 0$ is a linear relation.

A set which is not linearly independent is called *linearly dependent*.

If V is the space F^m and if the vectors $(v_1, \ldots, v_n)$ are given explicitly, we can decide linear independence by solving a system of homogeneous linear equations. For to say that a linear combination $x_1v_1 + \cdots + x_nv_n$ is zero means that each coordinate is zero, and this leads to m equations in the n unknowns x_i. For example, consider the set of three vectors

$$v_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad v_2 = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \quad v_3 = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}. \tag{3.5}$$

Let A denote the matrix whose columns are these vectors:

$$A = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 2 & 1 \\ 1 & 0 & 2 \end{bmatrix}. \tag{3.6}$$

A general linear combination of the vectors will have the form $x_1v_1 + x_2v_2 + x_3v_3$. Bringing the scalar coefficients to the other side, we can write this linear combination in the form AX, where $X = (x_1, x_2, x_3)^t$. Since $\det A = 1$, the equation $AX = 0$ has only the trivial solution, and this shows that (v_1, v_2, v_3) is a linearly independent set. On the other hand, if we add an arbitrary fourth vector v_4 to this set, the result will be linearly dependent, because every system of three homogeneous equations in four unknowns has a nontrivial solution [Chapter 1 (2.17)].

Here are some elementary facts about linear independence.

(3.7) **Proposition.**

(a) Any reordering of a linearly independent set is linearly independent.

(b) If $v_1 \in V$ is a nonzero vector, then the set (v_1) is linearly independent.

(c) A set (v_1, v_2) of two vectors is linearly dependent if and only if either $v_1 = 0$, or else v_2 is a multiple of v_1.

Let us verify the third of these assertions: Assume (v_1, v_2) dependent. Let the relation be $c_1v_1 + c_2v_2 = 0$, where c_1, c_2 are not both zero. If $c_2 \neq 0$, we can solve for v_2:

$$v_2 = \frac{-c_1}{c_2}v_1.$$

In this case v_2 is a multiple of v_1. If $c_2 = 0$, then $c_1 \neq 0$ and the equation shows that $v_1 = 0$. Conversely, if $v_2 = cv_1$, then the relation $cv_1 - v_2 = 0$ shows that the set (v_1, v_2) is linearly dependent, and if $v_1 = 0$, then the relation $v_1 + 0v_2 = 0$ shows the same thing. □

A set of vectors $(v_1,\dots,v_n)$ which is linearly independent and which also spans V is called a *basis*. For example, the vectors (1.4) form a basis for the space of solutions of the linear equation (1.3). We will often use a symbol such as **B** to denote a basis.

Let $\mathbf{B} = (v_1,\dots,v_n)$ be a basis. Then since **B** spans V, every $w \in V$ can be written as a linear combination (3.1). Since **B** is linearly independent, this expression is unique.

(3.8) **Proposition.** The set $\mathbf{B} = (v_1,\dots,v_n)$ is a basis if and only if every vector $w \in V$ can be written in a *unique* way in the form (3.1).

Proof. Suppose that **B** is a basis and that w is written as a linear combination in two ways, say (3.1) and also $w = c_1'v_1 + \cdots + c_1'v_n$. Then

$$0 = w - w = (c_1 - c_1')v_1 + \cdots + (c_n - c_n')v_n.$$

Hence by (3.4) $c_1 - c_1' = 0,\dots,c_n - c_n' = 0$. Thus the two linear combinations are the same. On the other hand, the definition of linear independence for **B** can be restated by saying that 0 has only one expression as a linear combination. This proves the converse. □

(3.9) **Example.** Let $V = F^n$ be the space of column vectors, and let e_i denote the column vector with 1 in the ith position and zeros elsewhere. The n vectors e_i form a basis for F^n called the *standard basis*. This basis was introduced before, in Chapter 1, Section 4. We will denote it by **E**. Every vector $X = (x_1,\dots,x_n)^t$ has the unique expression

$$X = x_1e_1 + \cdots + x_ne_n$$

as a linear combination of $\mathbf{E} = (e_1,\dots,e_n)$.

The set (3.5) is another basis of $\mathbb{R}^3$.

We now discuss the main facts (3.15–3.17) which relate the three notions of span, linear independence, and basis.

(3.10) **Proposition.** Let L be a linearly independent ordered set in V, and let $v \in V$ be any vector. Then the ordered set $L' = (L, v)$ obtained by adding v to L is linearly independent if and only if v is not in the subspace spanned by L.

Proof. Say that $L = (v_1,\dots,v_r)$. If $v \in \text{Span } L$, then $v = c_1v_1 + \cdots + c_rv_r$ for some $c_i \in F$. Hence

$$c_1v_1 + \cdots + c_rv_r + (-1)v = 0$$

is a linear relation among the vectors of L', and the coefficient -1 is not zero. Thus L' is linearly dependent.

Conversely, suppose that L' is linearly dependent, so that there is some linear relation

$$c_1v_1 + \cdots + c_rv_r + bv = 0,$$

in which not all coefficients are zero. Then certainly $b \neq 0$. For, if b were zero, the expression would reduce to

$$c_1v_1 + \cdots + c_rv_r = 0.$$

Since L is assumed to be linearly independent, we could conclude that $c_1 = \cdots = c_r = 0$ too, contrary to hypothesis. Now that we know $b \neq 0$, we can solve for v:

$$v = \frac{-c_1}{b}v_1 + \cdots + \frac{-c_r}{b}v_r.$$

Thus $v \in \text{Span } L$. □

(3.11) **Proposition.** Let S be an ordered set of vectors, let $v \in V$ be any vector, and let $S' = (S,v)$. Then Span S = Span S' if and only if $v \in$ Span S.

Proof. By definition, $v \in$ Span S'. So if $v \notin$ Span S, then Span $S \neq$ Span S'. Conversely, if $v \in$ Span S, then $S' \subset$ Span S; hence Span $S' \subset$ Span S (3.2). The fact that Span $S' \supset$ Span S is trivial, and so Span S' = Span S. □

(3.12) **Definition.** A vector space V is called *finite-dimensional* if there is some finite set S which spans V.

For the rest of this section, we assume that our given vector space V is finite-dimensional.

(3.13) **Proposition.** Any finite set S which spans V contains a basis. In particular, any finite-dimensional vector space has a basis.

Proof. Suppose $S = (v_1,\ldots,v_n)$ and that S is not linearly independent. Then there is a linear relation

$$c_1v_1 + \cdots + c_nv_n = 0$$

in which some c_i is not zero, say $c_n \neq 0$. Then we may solve for v_n:

$$v_n = \frac{-c_1}{c_n}v_1 + \cdots + \frac{-c_{n-1}}{c_n}v_{n-1}.$$

This shows that $v_n \in \text{Span}(v_1,\ldots,v_{n-1})$. Putting $v = v_n$ and $S = (v_1,\ldots,v_{n-1})$ in (3.11), we conclude $\text{Span}(v_1,\ldots,v_{n-1}) = \text{Span}(v_1,\ldots,v_n) = V$. So we may eliminate v_n from S. Continuing this way we eventually obtain a family which is linearly independent but still spans V—a basis.

Note. There is a problem with this proof if V is the zero vector space $\{0\}$. For, starting with an arbitrary collection of vectors in V (all of them equal to zero), our procedure will throw them out, one at a time, until there is only one vector $v_1 = 0$ left. And (0) is a linearly dependent set. How can we eliminate it? Of course the zero vector space is not particularly interesting. But it may lurk around, waiting to trip us up. We have to allow the possibility that a vector space which arises in the course of some computation, such as solving a system of homogeneous linear equations, is the zero space. In order to avoid having to make special mention of this case in the future, we adopt the following conventions:

(3.14) (a) The empty set is linearly independent.
(b) The span of the empty set is the zero subspace.

Thus the empty set is a basis for the zero vector space. These conventions allow us to throw out the last vector $v_1 = 0$, and rescue the proof. □

(3.15) **Proposition.** Let V be a finite-dimensional vector space. Any linearly independent set L can be extended by adding elements, to get a basis.

Proof. Let S be a finite set which spans V. If all elements of S are in Span L, then L spans V (3.2) and so it is a basis. If not, choose $v \in S$, which is not in Span L. By (3.10), (L, v) is linearly independent. Continue until you get a basis. □

(3.16) **Proposition.** Let S, L be finite subsets of V. Assume that S spans V and that L is linearly independent. Then S contains at least as many elements as L does.

Proof. To prove this, we write out what a relation of linear dependence on L means in terms of the set S, obtaining a homogeneous system of m linear equations in n unknowns, where $m = |S|$ and $n = |\mathrm{L}|$. Say that $S = (v_1, \ldots, v_m)$ and $L = (w_1, \ldots, w_n)$. We write each vector w_j as a linear combination of S, which we can do because S spans V, say

$$w_j = a_{1j}v_1 + \cdots + a_{mj}v_m = \sum_i a_{ij}v_i.$$

Let $u = c_1w_1 + \cdots + c_nw_n = \Sigma_j c_j w_j$ be a linear combination. Substituting, we obtain

$$u = \sum_{i,j} c_j a_{ij} v_i.$$

The coefficient of v_i in this sum is $\Sigma_j a_{ij}c_j$. If this coefficient is zero for every i, then $u = 0$. So to find a linear relation among the vectors of L, it suffices to solve the system $\Sigma_j a_{ij}x_j = 0$ of m equations in n unknowns. If $m < n$, then this system has a nontrivial solution [see Chapter 1 (2.17)], and therefore L is linearly dependent. □

(3.17) **Proposition.** Two bases $\mathbf{B}_1, \mathbf{B}_2$ of the vector space V have the same number of elements.

Proof. Put $\mathbf{B}_1 = S$, $\mathbf{B}_2 = L$ in (3.16) to get $|\mathbf{B}_1| \geq |\mathbf{B}_2|$. By symmetry, $|\mathbf{B}_2| \geq |\mathbf{B}_1|$. □

(3.18) **Definition.** The *dimension* of a finite-dimensional vector space V is the number of vectors in a basis. The dimension will be denoted by $\dim V$.

(3.19) **Proposition.**

(a) If S spans V, then $|S| \geq \dim V$, and equality holds only if S is a basis.

(b) If L is linearly independent, then $|L| \leq \dim V$, and equality holds only if L is a basis.

Proof. This follows from (3.13) and (3.15). □

(3.20) **Proposition.** If $W \subset V$ is a subspace of a finite-dimensional vector space, then W is finite-dimensional, and $\dim W \leq \dim V$. Moreover, $\dim W = \dim V$ only if $W = V$.

Proof. This will be obvious, once we show that W is finite-dimensional. For, if $W < V$, that is, if W is contained in but not equal to V, then a basis for W will not span V, but it can be extended to a basis of V by (3.15). Hence $\dim W < \dim V$. We now check finite-dimensionality: If some given linearly independent set L in W does not span W, there is a vector $w \in W$ not in Span L, and by Proposition (3.10), (L, w) is linearly independent. So, we can start with the empty set and add elements of W using (3.10), hoping to end up with a basis of W. Now it is obvious that if L is a linearly independent set in W then it is also linearly independent when viewed as a subset of V. Therefore (3.16) tells us that $|L| \leq n = \dim V$. So the process of adding vectors to L must come to an end after at most n steps. When it is impossible to apply (3.10) again, L is a basis of W. This shows that W is finite-dimensional, as required. □

Notes.

(a) The key facts to remember are (3.13), (3.15), and (3.16). The others follow.

(b) This material is not deep. Given the definitions, you could produce a proof of the main result (3.16) in a few days or less, though your first try would probably be clumsy.

One important example of a vector space is obtained from an arbitrary set S by forming linear combinations of elements of S with coefficients in F in a formal way. If $S = (s_1, \ldots, s_n)$ is a finite ordered set whose elements are distinct, then this space $V = V(S)$ is the set of all expressions

$$a_1 s_1 + \cdots + a_n s_n, \quad a_i \in F. \tag{3.21}$$

Addition and scalar multiplication are carried out formally, assuming no relations among the elements s_i:

(3.22)
$$(a_1s_1 + \cdots + a_ns_n) + (b_1s_1 + \cdots + b_ns_n) = (a_1 + b_1)s_1 + \cdots + (a_n + b_n)s_n$$
$$c(a_1s_1 + \cdots + a_ns_n) = (ca_1)s_1 + \cdots + (ca_n)s_n.$$

This vector space is isomorphic to F^n, by the correspondence

(3.23) $$(a_1, \ldots, a_n) \rightsquigarrow a_1s_1 + \cdots + a_ns_n.$$

Therefore the elements s_i, interpreted as the linear combinations

$$s_1 = 1s_1 + 0s_2 + \cdots + 0s_n,$$

form a basis which corresponds to the standard basis of F^n under the isomorphism (3.23). Because of this, $V(S)$ is often referred to as *the space with basis S,* or *the space of formal linear combinations of S.* If S is an infinite set, $V(S)$ is defined to be the space of all finite expressions (3.21), where $s_i \in S$ (see Section 5).

Since $V(S)$ is isomorphic to F^n when S contains n elements, there is no compelling logical reason for introducing it. However, in many applications, $V(S)$ has a natural interpretation. For example, if S is a set of ingredients, then a vector v may represent a recipe. Or if S is a set of points in the plane, then v (3.21) can be interpreted as a set of weights at the points of S.

4. COMPUTATION WITH BASES

The purpose of bases in vector spaces is to provide a method of computation, and we are going to learn to use them in this section. We will consider two topics: how to express a vector in terms of a given basis, and how to relate two different bases of the same vector space.

Suppose we are given a basis $(v_1, \ldots, v_n)$ of a vector space V. Remember: This means that every vector $v \in V$ can be expressed as a linear combination

(4.1) $$v = x_1v_1 + \cdots + x_nv_n, \quad x_i \in F,$$

in exactly one way. The scalars x_i are called the *coordinates* of v, and the column vector

(4.2) $$X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

is called the *coordinate vector* of v, with respect to the basis. We pose the problem of computing this coordinate vector.

The simplest case to understand is that V is the space of column vectors F^n.

Let $\mathbf{B} = (v_1,\dots,v_n)$ be a basis of F^n. Then each element v_i of our basis is a column vector, and so the array $(v_1,\dots,v_n)$ forms an $n \times n$ matrix. It seems advisable to introduce a new symbol for this matrix, so we will write it as

$$[\mathbf{B}] = \begin{bmatrix} | & & | \\ v_1 & \cdots & v_n \\ | & & | \end{bmatrix}. \tag{4.3}$$

For example, if $\mathbf{B}$ is the basis

$$v_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad v_2 = \begin{bmatrix} 3 \\ 5 \end{bmatrix}, \quad \text{then} \quad [\mathbf{B}] = \begin{bmatrix} 1 & 3 \\ 2 & 5 \end{bmatrix}. \tag{4.4}$$

If $\mathbf{E} = (e_1,\dots,e_n)$ is the standard basis, the matrix $[\mathbf{E}]$ is the identity matrix.

A linear combination $x_1v_1 + \cdots + x_nv_n$ can be written as the matrix product

$$[\mathbf{B}]X = \begin{bmatrix} | & & | \\ v_1 & \cdots & v_n \\ | & & | \end{bmatrix}\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = v_1x_1 + \cdots + v_nx_n, \tag{4.5}$$

where X denotes the column vector $(x_1,\dots,x_n)^t$. This is another example of block multiplication. The only new feature is that the definition of matrix multiplication has caused the scalar coefficients x_i to migrate to the right side of the vectors, which doesn't matter.

Now if a vector $Y = (y_1,\dots,y_n)^t$ is given, we can determine its coordinate vector with respect to the basis $\mathbf{B}$ by solving the equation

$$\begin{bmatrix} | & & | \\ v_1 & \cdots & v_n \\ | & & | \end{bmatrix}\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \quad \text{or} \quad [\mathbf{B}]X = Y \tag{4.6}$$

for the unknown vector X. This is done by inverting the matrix $[\mathbf{B}]$.

(4.7) **Proposition.** Let $\mathbf{B} = (v_1,\dots,v_n)$ be a basis of F^n, and let $Y \in F^n$ be a vector. The coordinate vector of Y with respect to the basis $\mathbf{B}$ is

$$X = [\mathbf{B}]^{-1}Y. \quad \square$$

Note that we get Y back if $\mathbf{B}$ is the standard basis $\mathbf{E}$, because $[\mathbf{E}]$ is the identity matrix. This is as it should be.

In Example (4.4),

$$[\mathbf{B}]^{-1} = \begin{bmatrix} 1 & 3 \\ 2 & 5 \end{bmatrix}^{-1} = \begin{bmatrix} -5 & 3 \\ 2 & -1 \end{bmatrix}.$$

So the coordinate vector of $Y = \begin{bmatrix} 1 \\ 4 \end{bmatrix}$ is $X = \begin{bmatrix} 7 \\ -2 \end{bmatrix}$, which means that $Y = 7v_1 - 2v_2$.

Of course we can not solve in this way unless the matrix is invertible. Fortunately, [**B**] is always invertible, and in fact it can be any invertible matrix.

(4.8) **Proposition.** Let A be an $n \times n$ matrix with entries in a field F. The columns of A form a basis of F^n if and only if A is invertible.

Proof. Denote the ith column of A by v_i. For any column vector $X = (x_1,\dots,x_n)^t$, the matrix product $AX = v_1x_1 + \cdots + v_nx_n$ is a linear combination of the set $(v_1,\dots,v_n)$. So this set is linearly independent if and only if the only solution of the equation $AX = 0$ is the trivial solution $X = 0$. And as we know, this is true if and only if A is invertible [Chapter 1 (2.18)]. Morever, if $(v_1,\dots,v_n)$ is a linearly independent set, then it forms a basis because the dimension of F^n is n. □

Now let V be an abstractly given vector space. We want to use matrix notation to facilitate the manipulation of bases, and the way we have written ordered sets of vectors was chosen with this in mind:

(4.9) $$(v_1,\dots,v_n).$$

Perhaps this array should be called a *hypervector*. Unless our vectors are given concretely, we won't be able to represent this hypervector by a matrix, so we will work with it formally, as if it were a vector. Since multiplication of two elements of a vector space is not defined, we can not multiply two matrices whose entries are vectors. But there is nothing to prevent us from multiplying the hypervector $(v_1,\dots,v_m)$ by a matrix of scalars. Thus a linear combination of these vectors can be written as the product with a column vector X:

(4.10) $$(v_1,\dots,v_m)\begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix} = v_1x_1 + \cdots + v_mx_m.$$

Evaluating the product, we obtain another vector—a linear combination. The scalar coefficients x_i are on the right side of the vectors as before. If we use a symbol such as **B** to denote the set $(v_1,\dots,v_m)$, then the notation for this linear combination becomes very compact: $\mathbf{B}X = v_1x_1 + \cdots + v_nx_n$.

We may also multiply a hypervector on the right by a matrix of scalars. If A is an $m \times n$ matrix, the product will be another hypervector, say $(w_1,\dots,w_n)$:

(4.11) $$(v_1,\dots,v_m)\begin{bmatrix} & A & \end{bmatrix} = (w_1,\dots,w_n).$$

To evaluate the product, we use the rule for matrix multiplication:

(4.12) $$w_j = v_1a_{1j} + v_2a_{2j} + \cdots + v_ma_{mj}.$$

So each vector w_j is a linear combination of $(v_1,\dots,v_m)$, and the scalar coefficients in

this linear combination form the columns of the matrix A. That is what the equation means. For example,

$$(v_1, v_2)\begin{bmatrix} 3 & 2 & 1 \\ 4 & 0 & 1 \end{bmatrix} = (3v_1 + 4v_2, 2v_1, v_1 + v_2).$$

Let us restate this formally:

(4.13) **Proposition.** Let $S = (v_1, \ldots, v_m)$ and $U = (w_1, \ldots, w_n)$ be ordered sets of elements of a vector space V. The elements of U are in the span of S if and only if there is an $m \times n$ scalar matrix A such that $(v_1, \ldots, v_m)A = (w_1, \ldots, w_n)$. □

Now let us consider the problem of determining the coordinate vector X of a given vector $v \in V$ with respect to a given basis $\mathbf{B} = (v_1, \ldots, v_n)$. That is, we wish to write $v = \mathbf{B}X$ explicitly, as in (4.10). It is clear that this is not possible unless both the basis and the vector are given in some explicit way, so we can not solve the problem as posed. But we can use multiplication by the hypervector $\mathbf{B}$ to define abstractly an *isomorphism of vector spaces*

(4.14) $$\psi\colon F^n \longrightarrow V \text{ sending}$$

$$X \rightsquigarrow \mathbf{B}X,$$

from the space F^n of column vectors to V. This map is bijective because every vector v is a linear combination (4.10) in exactly one way—it is surjective because the set $\mathbf{B}$ spans V, and injective because $\mathbf{B}$ is linearly independent. The axioms for an isomorphism (2.13) are easy to check. We can use this isomorphism to introduce *coordinates* into the vector space V.

The coordinate vector of a vector v is $X = \psi^{-1}(v)$. Please note that the symbol $\mathbf{B}^{-1}$ is not defined. So unless the basis is given more specifically, we won't have an explicit formula for the inverse function ψ^{-1}. But the existence of the isomorphism ψ is of interest in itself:

(4.15) **Corollary.** Every vector space V of dimension n is isomorphic to the space F^n of column vectors. □

Notice that F^n is *not* isomorphic to F^m if $m \neq n$, because F^n has a basis of n elements, and the number of elements in a basis depends only on the vector space, not on the choice of a basis. Thus the finite-dimensional vector spaces V over a field F are completely classified by (4.15): Every V is isomorphic to F^n, for some uniquely determined integer n. It follows that we will know all about an arbitrary vector space if we study the basic examples of column vectors. This reduces any problem on vector spaces to the familiar algebra of column vectors, once a basis is given.

We now come to a very important computational method: *change of basis*. Identifying V with the isomorphic vector space F^n is useful when a natural basis is

presented to us, but not when the given basis is poorly suited to the problem at hand. In that case, we will want to change coordinates. So let us suppose that we are given two bases for the same vector space V, say $\mathbf{B} = (v_1,\dots,v_n)$ and $\mathbf{B}' = (v_1',\dots,v_n')$. We will think of $\mathbf{B}$ as the *old* basis, and $\mathbf{B}'$ as a *new* basis. There are two computations which we wish to clarify. We ask first: How are the two bases related? Secondly, a vector $v \in V$ will have coordinates with respect to each of these bases, but of course they will be different. So we ask: How are the two coordinate vectors related? These are the computations called change of basis. They will be very important in later chapters. They are also confusing and can drive you nuts if you don't organize the notation well.

We begin by noting that since the new basis spans V, every vector of the old basis $\mathbf{B}$ is a linear combination of the new basis $\mathbf{B}' = (v_1',\dots,v_n')$. So Proposition (4.13) tells us that there is an equation of the form

$$(4.16) \qquad (v_1',\dots,v_n')\begin{bmatrix} & P & \end{bmatrix} = (v_1,\dots,v_n), \quad \text{or } \mathbf{B}'P = \mathbf{B},$$

where P is an $n \times n$ matrix of scalars. This matrix equation reads

$$(4.17) \qquad v_1' p_{1j} + v_2' p_{2j} + \cdots + v_n' p_{nj} = v_j,$$

where p_{ij} are the entries of P. The matrix P is called the *matrix of change of basis*. Its jth column is the coordinate vector of the old basis vector v_j, when computed with respect to the new basis $\mathbf{B}'$.

Note that the matrix of change of basis is *invertible*. This can be shown as follows: Interchanging the roles of $\mathbf{B}$ and $\mathbf{B}'$ provides a matrix P' such that $\mathbf{B}P' = \mathbf{B}'$. Combining this with (4.16), we obtain the relation $\mathbf{B}P'P = \mathbf{B}$:

$$(v_1,\dots,v_n)\begin{bmatrix} & P'P & \end{bmatrix} = (v_1,\dots,v_n).$$

This formula expresses each v_i as a linear combination of the vectors $(v_1,\dots v_n)$. The entries of the product matrix $P'P$ are the coefficients. But since $\mathbf{B}$ is a linearly independent set, there is *only one way* to write v_i as such a linear combination of $(v_1,\dots,v_n)$, namely $v_i = v_i$, or $\mathbf{B}I = \mathbf{B}$. So $P'P = I$. This shows that P is invertible.

Now let X be the coordinate vector of v, computed with respect to the old basis $\mathbf{B}$, that is, $v = \mathbf{B}X$. Substituting (4.16) gives us the matrix equation

$$(4.18) \qquad v = \mathbf{B}X = \mathbf{B}'PX.$$

This equation shows that $PX = X'$ is the coordinate vector of v with respect to the new basis $\mathbf{B}'$.

Recapitulating, we have a single matrix P, the matrix of change of basis, with the dual properties

$$(4.19) \qquad \mathbf{B} = \mathbf{B}'P \quad \text{and} \quad PX = X',$$

where X, X' denote the coordinate vectors of an arbitrary vector v with respect to the

two bases. Each of these properties characterizes P. Note the position of the primes carefully.

We can compute the matrix of change of basis explicitly when $V = F^n$ and the old basis is the standard basis $\mathbf{E}$, but where the new basis $\mathbf{B}'$ is arbitrary. The two bases determine matrices $[\mathbf{E}] = I$ and $[\mathbf{B}']$, as in (4.3). Formula (4.19) gives us the matrix equation $I = [\mathbf{B}']P$. Hence the matrix of change of basis is

(4.20) $$P = [\mathbf{B}']^{-1}, \quad \textit{if } V = F^n \textit{ and if the old basis is } \mathbf{E}.$$

We can also write this as $[\mathbf{B}'] = P^{-1}$. So

(4.21) *If the old basis is* $\mathbf{E}$, *the new basis vectors are the columns of* P^{-1}.

In the above discussion, the matrix P was determined in terms of two bases $\mathbf{B}$ and $\mathbf{B}'$. We could also turn the discussion around, starting with just one basis $\mathbf{B}$ and an invertible matrix $P \in GL_n(F)$. Then we can define a new basis by formula (4.16), that is,

(4.22) $$\mathbf{B}' = \mathbf{B}P^{-1}.$$

The vectors v_i making up the old basis are in the span of $\mathbf{B}'$ because $\mathbf{B} = \mathbf{B}'P$ (4.13). Hence $\mathbf{B}'$ spans V and, having the right number of elements, $\mathbf{B}'$ is a basis.

(4.23) **Corollary.** Let $\mathbf{B}$ be a basis of a vector space V. The other bases are the sets of the form $\mathbf{B}' = \mathbf{B}P^{-1}$, where $P \in GL_n(F)$ is an invertible matrix.

It is, of course, unnecessary to put an inverse matrix into this statement. Since P is arbitrary, so is P^{-1}. We could just as well set $P^{-1} = Q$ and say $\mathbf{B}' = \mathbf{B}Q$, where $Q \in GL_n(F)$. □

As an application of our discussion, let us compute the order of the general linear group $GL_2(F)$ when F is the prime field $\mathbb{F}_p$. We do this by computing the number of bases of the vector space $V = F^2$. Since the dimension of V is 2, any linearly independent set (v_1, v_2) of two elements forms a basis. The first vector v_1 of a linearly independent set is not zero. And since the order of F is p, V contains p^2 vectors including 0. So there are $p^2 - 1$ choices for the vector v_1. Next, a set (v_1, v_2) of two vectors, with v_1 nonzero, is linearly independent if and only if v_2 is not a multiple of v_1 (3.7). There are p multiples of a given nonzero vector v_1. Therefore if v_1 is given, there are $p^2 - p$ vectors v_2 such that (v_1, v_2) is linearly independent. This gives us

$$(p^2 - 1)(p^2 - p) = p(p + 1)(p - 1)^2$$

bases for V altogether.

(4.24) **Corollary.** The general linear group $GL_2(\mathbb{F}_p)$ has order $p(p + 1)(p - 1)^2$.

Proof. Proposition (4.23) establishes a bijective correspondence between bases of F^n and elements of $GL_n(F)$. □

5. INFINITE-DIMENSIONAL SPACES

Some vector spaces are too big to be spanned by any finite set of vectors. They are called *infinite-dimensional*. We are not going to need them very often, but since they are so important in analysis, we will discuss them briefly.

The most obvious example of an infinite-dimensional space is the space $\mathbb{R}^\infty$ of infinite real vectors

$$(a) = (a_1, a_2, a_3, \ldots). \tag{5.1}$$

It can also be thought of as the space of sequences $\{a_n\}$ of real numbers. Examples (1.7c, d) are also infinite-dimensional.

The space $\mathbb{R}^\infty$ has many important subspaces. Here are a few examples:

(5.2) **Examples.**

(a) Convergent sequences: $C = \{(a) \in \mathbb{R}^\infty \mid \lim_{n\to\infty} a_n \text{ exists}\}$.

(b) Bounded sequences: $\ell^\infty = \{(a) \in \mathbb{R}^\infty \mid \{a_n\} \text{ is bounded}\}$.

A sequence $\{a_n\}$ is called *bounded* if there is some real number b, a *bound*, such that $|a_n| \leq b$ for all n.

(c) Absolutely convergent series: $\ell^1 = \{(a) \in \mathbb{R}^\infty \mid \sum_1^\infty |a_n| < \infty\}$.

(d) Sequences with finitely many nonzero terms:

$$Z = \{(a) \in \mathbb{R}^\infty \mid a_n = 0 \text{ for all but finitely many } n\}.$$

All of the above subspaces are infinite-dimensional. You should be able to make up some more.

Now suppose that V is a vector space, infinite-dimensional or not. What should we mean by the *span* of an infinite set S of vectors? The difficulty is this: It is not always possible to assign a vector as the value of an infinite linear combination $c_1v_1 + c_2v_2 + \cdots$ in a consistent way. If we are talking about the vector space of real numbers, that is, $v_i \in \mathbb{R}^1$, then a value can be assigned provided that the series $c_1v_1 + c_2v_2 + \cdots$ converges. The same can be done for convergent series of vectors in $\mathbb{R}^n$ or $\mathbb{R}^\infty$. But many series don't converge, and then we don't know what value to assign.

In algebra it is customary to speak only of linear combinations of finitely many vectors. Therefore, the span of an infinite set S must be interpreted as the set of those vectors v which are linear combinations of *finitely many* elements of S:

$$v = c_1v_1 + \cdots + c_rv_r, \quad \text{where} \quad v_1, \ldots, v_r \in S. \tag{5.3}$$

The number r is allowed to be arbitrarily large, depending on the vector v:

$$\text{Span } S = \left\{\begin{matrix}\text{finite linear combinations} \\ \text{of elements of } S\end{matrix}\right\}. \tag{5.4}$$

With this definition, Propositions (3.2) and (3.11) continue to hold.

For example, let $e_i = (0,\dots,0,1,0,\dots)$ be the vector in $\mathbb{R}^\infty$ with 1 in the ith position as its only nonzero coordinate. Let $S = (e_1, e_2, e_3, \dots)$ be the infinite set of these vectors e_i. The set S does not span $\mathbb{R}^\infty$, because the vector

$$w = (1, 1, 1, \dots)$$

is not a (finite) linear combination. Instead the span of S is the subspace Z (5.2d).

A set S, infinite or not, is called *linearly independent* if there is no *finite* relation

$$c_1 v_1 + \cdots + c_r v_r = 0, \quad v_1, \dots, v_r \in S, \tag{5.5}$$

except for the trivial relation, in which $c_1 = \cdots = c_r = 0$. Again, the number r is allowed to be arbitrary, that is, the condition has to hold for arbitrarily large r and arbitrary vectors $v_1, \dots, v_r \in S$. For example, the set $S' = (w; e_1, e_2, e_3, \dots)$ is linearly independent, if w, e_i are the vectors defined as above. With this definition of linear independence, Proposition (3.10) continues to be true.

As with finite sets, a *basis* S of V is a linearly independent set which spans V. Thus $S = (e_1, e_2, \dots)$ is a basis of the space Z. It can be shown, using the *Axiom of Choice*, that every vector space V has a basis. However, the proof doesn't tell you how to get one. A basis for $\mathbb{R}^\infty$ will have uncountably many elements, and therefore it can not be written down in an explicit way. We won't need bases for infinite-dimensional spaces very often.

Let us go back for a moment to the case that our vector space V is finite-dimensional (3.12), and ask if there can be an *infinite* basis. In Section 3, we saw that any two finite bases have the same number of elements. We will now complete the picture by showing that every basis is finite. The only confusing point is taken care of by the following proposition:

(5.6) **Proposition.** Let V be finite-dimensional, and let S be any set which spans V. Then S contains a finite subset which spans V.

Proof. By assumption, there is some finite set, say $(w_1, \dots, w_m)$, which spans V. Each w_i is a linear combination of finitely many elements of S, since Span $S = V$. So when we express the vectors $w_1, \dots, w_m$ in terms of the set S, we only need to use finitely many of its elements. The ones we use make up a finite subset $S' \subset S$. So, $(w_1, \dots, w_m) \subset \text{Span } S'$. Since $(w_1, \dots, w_m)$ spans V, so does S'. □

(5.7) **Proposition.** Let V be a finite-dimensional vector space.

(a) Every set S which spans V contains a finite basis.

(b) Every linearly independent set L is finite and therefore extends to a finite basis.

(c) Every basis is finite.

We leave the proof of (5.7) as an exercise. □

6. DIRECT SUMS

Let V be a vector space, and let $W_1,\dots,W_n$ be subspaces of V. Much of the treatment of linear independence and spans of vectors has analogues for subspaces, and we are going to work out these analogues here.

We consider vectors $v \in V$ which can be written as a sum

$$v = w_1 + \cdots + w_n, \tag{6.1}$$

where w_i is a vector in W_i. The set of all such vectors is called the *sum* of the subspaces or their *span,* and is denoted by

$$W_1 + \cdots + W_n = \{v \in V \mid v = w_1 + \cdots + w_n, \quad \text{with } w_i \in W_i\}. \tag{6.2}$$

The sum is a subspace of V, analogous to the span of a set $\{v_1,\dots,v_n\}$ of vectors. Clearly, it is the smallest subspace containing $W_1,\dots,W_n$.

The subspaces $W_1,\dots,W_n$ are called *independent* if no sum $w_1 + \cdots + w_n$ with $w_i \in W$ is zero, except for the trivial sum in which $w_i = 0$ for all i. In other words, the spaces are independent if

$$w_1 + \cdots + w_n = 0 \text{ and } w_i \in W_i \text{ implies } w_i = 0 \text{ for all } i. \tag{6.3}$$

In case the span is the whole space and the subspaces are independent, we say that V is the *direct sum* of $W_1,\dots,W_n$, and we write

$$V = W_1 \oplus \cdots \oplus W_n, \text{ if } V = W_1 + \cdots + W_n \text{ and if } W_1,\dots,W_n \text{ are independent.} \tag{6.4}$$

This is equivalent to saying that every vector $v \in V$ can be written in the form (6.1) in *exactly one way.*

So, if $W_1,\dots,W_n$ are independent subspaces of a vector space V and if $U = W_1 + \cdots + W_n$ is their sum, then in fact U is their direct sum: $U = W_1 \oplus \cdots \oplus W_n$.

We leave the proof of the following two propositions as an exercise.

(6.5) **Proposition.**

(a) A single subspace W_1 is independent.

(b) Two subspaces W_1, W_2 are independent if and only if $W_1 \cap W_2 = (0)$. □

(6.6) **Proposition.** Let $W_1,\dots,W_n$ be subspaces of a finite-dimensional vector space V, and let $\mathbf{B}_i$ be a basis for W_i.

(a) The ordered set $\mathbf{B}$ obtained by listing the bases $\mathbf{B}_1,\dots,\mathbf{B}_n$ in order is a basis of V if and only if V is the direct sum $W_1 \oplus \cdots \oplus W_n$.

(b) $\dim(W_1 + \cdots + W_n) \leq (\dim W_1) + \cdots + (\dim W_n)$, with equality if and only if the spaces are independent. □

(6.7) **Corollary.** Let W be a subspace of a finite-dimensional vector space V. There is another subspace W' such that $V = W \oplus W'$.

Proof. Let $(w_1,\dots,w_d)$ be a basis for W. Extend to a basis $(w_1,\dots,w_d; v_1,\dots,v_{n-d})$ for V (3.15). The span of $(v_1,\dots,v_{n-d})$ is the required subspace W'. □

(6.8) **Example.** Let $v_1,\dots,v_n$ be nonzero vectors, and let W_i be the span of the single vector v_i. This is the one-dimensional subspace which consists of all scalar multiples of v_i: $W_i = \{cv_i\}$. Then $W_1,\dots,W_n$ are independent subspaces if and only if $(v_1,\dots,v_n)$ are independent vectors. This becomes clear if we compare (3.4) and (6.3). The statement in terms of subspaces is actually the neater one, because the scalar coefficients are absorbed.

(6.9) **Proposition.** Let W_1, W_2 be subspaces of a finite-dimensional vector space V. Then

$$\dim W_1 + \dim W_2 = \dim(W_1 \cap W_2) + \dim(W_1 + W_2).$$

Proof. Note first that the intersection of two subspaces is again a subspace. Choose a basis $(u_1,\dots,u_r)$ for the space $W_1 \cap W_2$, where $r = \dim(W_1 \cap W_2)$. This is a linearly independent set, and it is in W_1. Hence we can extend it to a basis of W_1, say

$$(u_1,\dots,u_r; x_1,\dots,x_{m-r}), \tag{6.10}$$

where $m = \dim W_1$. Similarly, we can extend it to a basis

$$(u_1,\dots,u_r; y_1,\dots,y_{n-r}), \tag{6.11}$$

of W_2, where $n = \dim W_2$. The proposition will follow if we show that the set

$$(u_1,\dots,u_r; x_1,\dots,x_{m-r}; y_1,\dots,y_{n-r}) \tag{6.12}$$

is a basis of $W_1 + W_2$.

This assertion has two parts. First, the vectors (6.12) span $W_1 + W_2$. For any vector v in $W_1 + W_2$ is a sum $v = w_1 + w_2$, with $w_i \in W_i$. We can write w_1 as a linear combination of (6.10), and w_2 as a linear combination of (6.11). Collecting terms, we find that v is a linear combination of (6.12).

Next, the vectors (6.11) are linearly independent: Suppose that some linear combination is zero, say

$$a_1u_1 + \cdots + a_ru_r + b_1x_1 + \cdots + b_{m-r}x_{m-r} + c_1y_1 + \cdots + c_{n-r}y_{n-r} = 0.$$

Abbreviate this as $u + x + y = 0$. Solve for y: $y = -u - x \in W_1$. But $y \in W_2$ too. Hence $y \in W_1 \cap W_2$, and so y is a linear combination, say u', of $(u_1,\dots,u_r)$. Then $-u' + y = 0$ is a relation among the vectors (6.11), which are independent. So it must be the trivial relation. This shows that $y = 0$. Thus our original relation reduces to $u + x = 0$. Since (6.10) is a basis, this relation is trivial: $u = 0$ and $x = 0$. So the whole relation was trivial, as required. □

I don't need to learn 8 + 7: I'll remember 8 + 8 and subtract 1.

T. Cuyler Young, Jr.

EXERCISES

1. Real Vector Spaces

1. Which of the following subsets of the vector space of real $n \times n$ matrices is a subspace?
   (a) symmetric matrices ($A = A^t$)
   (b) invertible matrices
   (c) upper triangular matrices
2. Prove that the intersection of two subspaces is a subspace.
3. Prove the cancellation law in a vector space: If $cv = cw$ and $c \neq 0$, then $v = w$.
4. Prove that if w is an element of a subspace W, then $-w \in W$ too.
5. Prove that the classification of subspaces of $\mathbb{R}^3$ stated after (1.2) is complete.
6. Prove that every solution of the equation $2x_1 - x_2 - 2x_3 = 0$ has the form (1.5).
7. What is the description analogous to (1.4) obtained from the particular solutions $u_1 = (2, 2, 1)$ and $u_2 = (0, 2, -1)$?

2. Abstract Fields

1. Prove that the set of numbers of the form $a + b\sqrt{2}$, where a, b are rational numbers, is a field.
2. Which subsets of $\mathbb{C}$ are closed under $+$, $-$, $\times$, and $\div$ but fail to contain 1?
3. Let F be a subset of $\mathbb{C}$ such that F^+ is a subgroup of $\mathbb{C}^+$ and $F^\times$ is a subgroup of $\mathbb{C}^\times$. Prove that F is a subfield of $\mathbb{C}$.
4. Let $V = F^n$ be the space of column vectors. Prove that every subspace W of V is the space of solutions of some system of homogeneous linear equations $AX = 0$.
5. Prove that a nonempty subset W of a vector space satisfies the conditions (2.12) for a subspace if and only if it is closed under addition and scalar multiplication.
6. Show that in Definition (2.3), axiom (ii) can be replaced by the following axiom: $F^\times$ is an abelian group, and $1 \neq 0$. What if the condition $1 \neq 0$ is omitted?
7. Define homomorphism of fields, and prove that every homomorphism of fields is injective.
8. Find the inverse of 5 (modulo p) for $p = 2, 3, 7, 11, 13$.
9. Compute the polynomial $(x^2 + 3x + 1)(x^3 + 4x^2 + 2x + 2)$ when the coefficients are regarded as elements of the fields (a) $\mathbb{F}_5$ (b) $\mathbb{F}_7$.
10. Consider the system of linear equations $\begin{bmatrix} 8 & 3 \\ 2 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ -1 \end{bmatrix}$.
    (a) Solve it in $\mathbb{F}_p$ when $p = 5, 11, 17$.
    (b) Determine the number of solutions when $p = 7$.

11. Find all primes p such that the matrix

$$A = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 3 & -1 \\ -2 & 0 & 2 \end{bmatrix}$$

is invertible, when its entries are considered to be in $\mathbb{F}_p$.

12. Solve completely the systems of linear equations $AX = B$, where

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & -1 & -1 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

(a) in $\mathbb{Q}$ **(b)** in $\mathbb{F}_2$ **(c)** in $\mathbb{F}_3$ **(d)** in $\mathbb{F}_7$.

13. Let p be a prime integer. The nonzero elements of $\mathbb{F}_p$ form a group $\mathbb{F}_p^\times$ of order $p - 1$. It is a fact that this group is always cyclic. Verify this for all primes $p < 20$ by exhibiting a generator.

14. **(a)** Let p be a prime. Use the fact that $\mathbb{F}_p^\times$ is a group to prove that $a^{p-1} \equiv 1$ (modulo p) for every integer a not congruent to zero.

(b) Prove *Fermat's Theorem:* For every integer a,

$$a^p \equiv a \text{ (modulo } p).$$

15. **(a)** By pairing elements with their inverses, prove that the product of all nonzero elements of $\mathbb{F}_p$ is -1.

(b) Let p be a prime integer. Prove *Wilson's Theorem:*

$$(p - 1)! \equiv -1 \text{ (modulo } p).$$

16. Consider a system $AX = B$ of n linear equations in n unknowns, where A and B have integer entries. Prove or disprove: If the system has an integer solution, then it has a solution in $\mathbb{F}_p$ for all p.

17. Interpreting matrix entries in the field $\mathbb{F}_2$, prove that the four matrices $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$ form a field.

18. The proof of Lemma (2.8) contains a more direct proof of (2.6). Extract it.

3. Bases and Dimension

1. Find a basis for the subspace of $\mathbb{R}^4$ spanned by the vectors $(1, 2, -1, 0)$, $(4, 8, -4, -3)$, $(0, 1, 3, 4)$, $(2, 5, 1, 4)$.

2. Let $W \subset \mathbb{R}^4$ be the space of solutions of the system of linear equations $AX = 0$, where $A = \begin{bmatrix} 2 & 1 & 2 & 3 \\ 1 & 1 & 3 & 0 \end{bmatrix}$. Find a basis for W.

3. **(a)** Show that a subset of a linearly independent set is linearly independent.

(b) Show that any reordering of a basis is also a basis.

4. Let V be a vector space of dimension n over F, and let $0 \leq r \leq n$. Prove that V contains a subspace of dimension r.

5. Find a basis for the space of symmetric $n \times n$ matrices.
6. Prove that a square matrix A is invertible if and only if its columns are linearly independent.
7. Let V be the vector space of functions on the interval $[0, 1]$. Prove that the functions x^3, $\sin x$, and $\cos x$ are linearly independent.
8. Let A be an $m \times n$ matrix, and let A' be the result of a sequence of elementary row operations on A. Prove that the rows of A span the same subspace as the rows of A'.
9. Let V be a complex vector space of dimension n. Prove that V has dimension $2n$ as real vector space.
10. A complex $n \times n$ matrix is called *hermitian* if $a_{ij} = \bar{a}_{ji}$ for all i, j. Show that the hermitian matrices form a real vector space, find a basis for that space, and determine its dimension.
11. How many elements are there in the vector space $\mathbb{F}_p{}^n$?
12. Let $F = \mathbb{F}_2$. Find all bases of F^2.
13. Let $F = \mathbb{F}_5$. How many subspaces of each dimension does the space F^3 contain?
14. **(a)** Let V be a vector space of dimension 3 over the field $\mathbb{F}_p$. How many subspaces of each dimension does V have?
    (b) Answer the same question for a vector space of dimension 4.
15. **(a)** Let $F = \mathbb{F}_2$. Prove that the group $GL_2(F)$ is isomorphic to the symmetric group S_3.
    (b) Let $F = \mathbb{F}_3$. Determine the orders of $GL_2(F)$ and of $SL_2(F)$.
16. Let W be a subspace of V.
    (a) Prove that there is a subspace U of V such that $U + W = V$ and $U \cap W = 0$.
    (b) Prove that there is no subspace U such that $W \cap U = 0$ and that $\dim W + \dim U > \dim V$.

4. Computation with Bases

1. Compute the matrix P of change of basis in F^2 relating the standard basis $\mathbf{E}$ to $\mathbf{B}' = (v_1, v_2)$, where $v_1 = (1, 3)^t$, $v_2 = (2, 2)^t$.
2. Determine the matrix of change of basis, when the old basis is the standard basis $(e_1, \ldots, e_n)$ and the new basis is $(e_n, e_{n-1}, \ldots, e_1)$.
3. Determine the matrix P of change of basis when the old basis is (e_1, e_2) and the new basis is $(e_1 + e_2, e_1 - e_2)$.
4. Consider the equilateral coordinate system for $\mathbb{R}^2$, given by the basis $\mathbf{B}'$ in which $v_1 = e_1$ and v_2 is a vector of unit length making an angle of $120°$ with v_1. Find the matrix relating the standard basis $\mathbf{E}$ to $\mathbf{B}'$.
5. (i) Prove that the set $\mathbf{B} = ((1, 2, 0)^t, (2, 1, 2)^t, (3, 1, 1)^t)$ is a basis of $\mathbb{R}^3$.
   (ii) Find the coordinate vector of the vector $v = (1, 2, 3)^t$ with respect to this basis.
   (iii) Let $\mathbf{B}' = ((0, 1, 0)^t, (1, 0, 1)^t, (2, 1, 0)^t)$. Find the matrix P relating $\mathbf{B}$ to $\mathbf{B}'$.
   (iv) For which primes p is $\mathbf{B}$ a basis of $\mathbb{F}_p{}^3$?
6. Let $\mathbf{B}$ and $\mathbf{B}'$ be two bases of the vector space F^n. Prove that the matrix of change of basis is $P = [\mathbf{B}']^{-1}[\mathbf{B}]$.
7. Let $\mathbf{B} = (v_1, \ldots, v_n)$ be a basis of a vector space V. Prove that one can get from $\mathbf{B}$ to any other basis $\mathbf{B}'$ by a finite sequence of steps of the following types:

(i) Replace v_i by $v_i + av_j$, $i \neq j$, for some $a \in F$.
(ii) Replace v_i by cv_i for some $c \neq 0$.
(iii) Interchange v_i and v_j.

8. Rewrite the proof of Proposition (3.16) using the notation of Proposition (4.13).
9. Let $V = F^n$. Establish a bijective correspondence between the sets $\mathcal{B}$ of bases of V and $GL_n(F)$.
10. Let F be a field containing 81 elements, and let V be a vector space of dimension 3 over F. Determine the number of one-dimensional subspaces of V.
11. Let $F = \mathbb{F}_p$.
    (a) Compute the order of $SL_2(F)$.
    (b) Compute the number of bases of F^n, and the orders of $GL_n(F)$ and $SL_n(F)$.
12. **(a)** Let A be an $m \times n$ matrix with $m < n$. Prove that A has no left inverse by comparing A to the square $n \times n$ matrix obtained by adding $(n - m)$ rows of zeros at the bottom.
    (b) Let $\mathbf{B} = (v_1, \ldots, v_m)$ and $\mathbf{B}' = (v_1', \ldots, v_n')$ be two bases of a vector space V. Prove that $m = n$ by defining matrices of change of basis and showing that they are invertible.

5. Infinite-Dimensional Spaces

1. Prove that the set $(w; e_1, e_2, \ldots)$ introduced in the text is linearly independent, and describe its span.
2. We could also consider the space of doubly infinite sequences $(a) = (\ldots, a_{-1}, a_0, a_1, \ldots)$, with $a_i \in \mathbb{R}$. Prove that this space is isomorphic to $\mathbb{R}^\infty$.
3. Prove that the space Z is isomorphic to the space of real polynomials.
4. Describe five more infinite-dimensional subspaces of the space $\mathbb{R}^\infty$.
5. For every positive integer, we can define the space ℓ^p to be the space of sequences such that $\Sigma |a_i|^p < \infty$.
    (a) Prove that ℓ^p is a subspace of $\mathbb{R}^\infty$.
    (b) Prove that $\ell^p < \ell^{p+1}$.
6. Let V be a vector space which is spanned by a countably infinite set. Prove that every linearly independent subset of V is finite or countably infinite.
7. Prove Proposition (5.7).

6. Direct Sums

1. Prove that the space $\mathbb{R}^{n \times n}$ of all $n \times n$ real matrices is the direct sum of the spaces of symmetric matrices ($A = A^t$) and of skew-symmetric matrices ($A = -A^t$).
2. Let W be the space of $n \times n$ matrices whose trace is zero. Find a subspace W' so that $\mathbb{R}^{n \times n} = W \oplus W'$.
3. Prove that the sum of subspaces is a subspace.
4. Prove Proposition (6.5).
5. Prove Proposition (6.6).

Miscellaneous Problems

1. **(a)** Prove that the set of symbols $\{a + bi \mid a, b \in \mathbb{F}_3\}$ forms a field with nine elements, if the laws of composition are made to mimic addition and multiplication of complex numbers.

(b) Will the same method work for $\mathbb{F}_5$? For $\mathbb{F}_7$? Explain.

***2.** Let V be a vector space over an infinite field F. Prove that V is not the union of finitely many proper subspaces.

***3.** Let W_1, W_2 be subspaces of a vector space V. The formula $\dim(W_1 + W_2) = \dim W_1 + \dim W_2 - \dim(W_1 \cap W_2)$ is analogous to the formula $|S_1 \cup S_2| = |S_1| + |S_2| - |S_1 \cap S_2|$, which holds for sets. If three sets are given, then

$$|S_1 \cup S_2 \cup S_3| = |S_1| + |S_2| + |S_3| - |S_1 \cap S_2| - |S_1 \cap S_3| - |S_2 \cap S_3| + |S_1 \cap S_2 \cap S_3|.$$

Does the corresponding formula for dimensions of subspaces hold?

4. Let F be a field which is not of characteristic 2, and let $x^2 + bx + c = 0$ be a quadratic equation with coefficients in F. Assume that the discriminant $b^2 - 4c$ is a square in F, that is, that there is an element $\delta \in F$ such that $\delta^2 = b^2 - 4c$. Prove that the quadratic formula $x = (-b + \delta)/2a$ solves the quadratic equation in F, and that if the discriminant is not a square the polynomial has no root in F.

5. **(a)** What are the orders of the elements $\begin{bmatrix} 1 & 1 \\ & 1 \end{bmatrix}, \begin{bmatrix} 2 & \\ & 1 \end{bmatrix}$ of $GL_2(\mathbb{R})$?

(b) Interpret the entries of these matrices as elements of $\mathbb{F}_7$, and compute their orders in the group $GL_2(\mathbb{F}_7)$.

6. Consider the function det: $F^{n\times n} \longrightarrow F$, where $F = \mathbb{F}_p$ is a finite field with p elements and $F^{n\times n}$ is the set of 2×2 matrices.

(a) Show that this map is surjective.

(b) Prove that all nonzero values of the determinant are taken on the same number of times.

7. Let A be an $n \times n$ real matrix. Prove that there is a polynomial $f(t) = a_r t^r + a_{r-1}t^{r-1} + \cdots + a_1 t + a_0$ which has A as root, that is, such that $a_r A^r + a_{r-1}A^{r-1} + \cdots + a_1 A + a_0 I = 0$. Do this by showing that the matrices $I, A, A^2, \ldots$ are linearly dependent.

***8.** An *algebraic curve* in $\mathbb{R}^2$ is the locus of zeros of a polynomial $f(x, y)$ in two variables. By a *polynomial path* in $\mathbb{R}^2$, we mean a parametrized path $x = x(t)$, $y = y(t)$, where $x(t)$, $y(t)$ are polynomials in t.

(a) Prove that every polynomial path lies on a real algebraic curve by showing that, for sufficiently large n, the functions $x(t)^i y(t)^j$, $0 \le i, j \le n$, are linearly dependent.

(b) Determine the algebraic curve which is the image of the path $x = t^2 + t$, $y = t^3$ explicitly, and draw it.