chapter I | Vector spaces

In this chapter we introduce the concepts of a vector space and a basis for that vector space. We assume that there is at least one basis with a finite number of elements, and this assumption enables us to prove that the vector space has a vast variety of different bases but that they all have the same number of elements. This common number is called the dimension of the vector space.

For each choice of a basis there is a one-to-one correspondence between the elements of the vector space and a set of objects we shall call n-tuples. A different choice for a basis will lead to a different correspondence between the vectors and the n-tuples. We regard the vectors as the fundamental objects under consideration and the n-tuples a representations of the vectors. Thus, how a particular vector is represented depends on the choice of the basis, and these representations are non-invariant. We call the n-tuple the coordinates of the vector it represents; each basis determines a coordinate system.

We then introduce the concept of subspace of a vector space and develop the algebra of subspaces. Under the assumption that the vector space is finite dimensional, we prove that each subspace has a basis and that for each basis of the subspace there is a basis of the vector space which includes the basis of the subspace as a subset.

1 | Definitions

To deal with the concepts that are introduced we adopt some notational conventions that are commonly used. We usually use sans-serif italic letter to denote sets.

$\alpha \in S$ means α is an *element* of the set S.
$\alpha \notin S$ means α is *not an element* of the set S.

$S \subset T$ means S is a *subset* of the set T.

$S \cap T$ denotes the *intersection* of the sets S and T, the set of elements in both S and T.

$S \cup T$ denotes the *union* of the sets S and T, the set of elements in S or T.

$T - S$ denotes the set of elements in T but not in S. In case T is the set of all objects under consideration, we shall call $T - S$ the *complement* of S and denote it by CS.

$S_\mu : \mu \in M$ denotes a collection of sets indexed so that one set S_μ is specified for each element $\mu \in M$. M is called the *index set*.

$\cap_{\mu \in M} S_\mu$ denotes the intersection of all sets $S_\mu : \mu \in M$.

$\cup_{\mu \in M} S_\mu$ denotes the union of all sets $S_\mu : \mu \in M$.

$\emptyset$ denotes the set with no elements, the *empty set*.

A set will often be specified by listing the elements in the set or by giving a property which characterizes the elements of the set. In such cases we use braces: $\{\alpha, \beta\}$ is the set containing just the elements α and β, $\{\alpha \mid \mathrm{P}\}$ is the set of all α with property P, $\{\alpha_\mu \mid \mu \in M\}$ denotes the set of all α_μ corresponding to μ in the index set M. We have such frequent use for the set of all integers or a subset of the set of all integers as an index set that we adopt a special convention for these cases. $\{\alpha_i\}$ denotes a set of elements indexed by a subset of the set of integers. Usually the same index set is used over and over. In such cases it is not necessary to repeat the specifications of the index set and often designation of the index set will be omitted. Where clarity requires it, the index set will be specified. We are careful to distinguish between the set $\{\alpha_i\}$ and an element α_i of that set.

Definition. By a *field* F we mean a non-empty set of elements with two laws of combination, which we call addition and multiplication, satisfying the following conditions:

*F*1. To every pair of elements $a, b \in F$ there is associated a unique element, called their sum, which we denote by $a + b$.

*F*2. Addition is associative; $(a + b) + c = a + (b + c)$.

*F*3. There exists an element, which we denote by 0, such that $a + 0 = a$ for all $a \in F$.

*F*4. For each $a \in F$ there exists an element, which we denote by $-a$, such that $a + (-a) = 0$. Following usual practice we write $b + (-a) = b - a$.

*F*5. Addition is commutative; $a + b = b + a$.

*F*6. To every pair of element $a, b \in F$ there is associated a unique element, called their product, which we denote by ab, or $a \cdot b$.

*F*7. Multiplication is associative; $(ab)c = a(bc)$.

*F*8. There exists an element different from 0, which we denote by 1, such that $a \cdot 1 = a$ for all $a \in F$.

*F*9. For each $a \in F$, $a \neq 0$, there exists an element which we denote by a^{-1}, such that $a \cdot a^{-1} = 1$.

*F*10. Multiplicatiôn is commutative: $ab = ba$.

*F*11. Multiplication is distributive with respect to addition:

$$(a + b)c = ac + bc.$$

The elements of F are called *scalars*, and will generally be denoted by lower case Latin italic letters.

The rational numbers, real numbers, and complex numbers are familiar and important examples of fields, but they do not exhaust the possibilities. As a less familiar example, consider the set $\{0, 1\}$ where addition is defined by the rules: $0 + 0 = 1 + 1 = 0, 0 + 1 = 1$; and multiplication is defined by the rules: $0 \cdot 0 = 0 \cdot 1 = 0$, $1 \cdot 1 = 1$. This field has but two elements, and there are other fields with finitely many elements.

We do not develop the various properties of abstract fields and we are not concerned with any specific field other than the rational numbers, the real numbers, and the complex numbers. We find it convenient and desirable at the moment to leave the exact nature of the field of scalars unspecified because much of the theory of vector spaces and matrices is valid for arbitrary fields.

The student unacquainted with the theory of abstract fields will not be handicapped for it will be sufficient to think of F as being one of the familiar fields. All that matters is that we can perform the operations of addition and subtraction, multiplication and division, in the usual way. Later we have to restrict F to either the field of real numbers or the field of complex numbers in order to obtain certain classical results; but we postpone that moment as long as we can. At another point we have to make a very mild assumption, that is, $1 + 1 \neq 0$, a condition that happens to be false in the example given above. The student interested mainly in the properties of matrices with real or complex coefficients should consider this to be no restriction.

Definition. A *vector space* V over F is a non-empty set of elements, called *vectors*, with two laws of combination, called *vector addition* (or *addition*) and *scalar multiplication*, satisfying the following conditions:

*A*1. To every pair of vectors $\alpha, \beta \in V$ there is associated a unique vector in V called their *sum*, which we denote by $\alpha + \beta$.

*A*2. Addition is associative; $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$.

*A*3. There exists a vector, which we denote by 0, such that $\alpha + 0 = \alpha$ for all $\alpha \in V$.

*A*4. For each $\alpha \in V$ there exists an element, which we denote by $-\alpha$, such that $\alpha + (-\alpha) = 0$.

*A*5. Addition is commutative; $\alpha + \beta = \beta + \alpha$.

$B1$. To every scalar $a \in F$ and vector $\alpha \in V$, there is associated a unique vector, called the *product* of a and α, which we denote by $a\alpha$.

$B2$. Scalar multiplication is associative: $a(b\alpha) = (ab)\alpha$.

$B3$. Scalar multiplication is distributive with respect to vector addition; $a(\alpha + \beta) = a\alpha + a\beta$.

$B4$. Scalar multiplication is distributive with respect to scalar addition; $(a + b)\alpha = a\alpha + b\alpha$.

$B5$. $1 \cdot \alpha = \alpha$ (where $1 \in F$).

We generally use lower case Greek letters to denote vectors. An exception is the zero vector of $A3$. From a logical point of view we should not use the same symbol "0 ' for both the zero scalar and the zero vector, but this practice is rooted in a long tradition and it is not as confusing as it may seem at first.

The vector space axioms concerning addition alone have already appeared in the definition of a field. A set of elements satisfying the first four axioms is called a *group*. If the set of elements also satisfies $A5$ it is called a *commutative* group or *abelian* group. Thus both fields and vector spaces are abelian groups under addition. The theory of groups is well developed and our subsequent discussion would be greatly simplified if we were to assume a prior knowledge of the theory of groups. We do not assume a prior knowledge of the theory of groups; therefore, we have to develop some of their elementary properties as we go along, although we do not stop to point out that what was proved is properly a part of group theory. Except for specific applications in Chapter VI we do no more than use the term "group" to denote a set of elements satisfying these conditions.

First, we give some examples of vector spaces. Any notation other than "F" for a field and "V" for a vector space is used consistently in the same way throughout the rest of the book, and these examples serve as definitions for these notations:

(1) Let F be any field and let $V = P$ be the set of all polynomials in an indeterminate x with coefficients in F. Vector addition is defined to be the ordinary addition of polynomials, and multiplication is defined to be the ordinary multiplication of a polynomial by an element of F.

(2) For any positive integer n, let P_n be the set of all polynomials in x with coefficients in F of degree $\leq n - 1$, together with the zero polynomial. The operations are defined as in Example (1).

(3) Let $F = R$, the field of real numbers, and take V to be the set of all real-valued functions of a real variable. If f and g are functions we define vector addition and scalar multiplication by the rules

$$\begin{aligned}(f + g)(x) &= f(x) + g(x),\\ (af)(x) &= a[f(x)].\end{aligned} \tag{1.1}$$

(4) Let $F = R$, and let V be the set of continuous real-valued functions of a real variable. The operations are defined as in Example (3). The point of this example is that it requires a theorem to show that $A1$ and $B1$ are satisfied.

(5) Let $F = R$, and let V be the set of real-valued functions defined on the interval $[0, 1]$ and integrable over that interval. The operations are defined as in Example (3). Again, the main point is to show that $A1$ and $B1$ are satisfied.

(6) Let $F = R$, and let V be the set of all real-valued functions of a real variable differentiable at least m times (m a positive integer). The operations are defined as in Example (3).

(7) Let $F = R$, and let V be the set of all real-valued functions differentiable at least twice and satisfying the differential equation $\frac{d^2y}{dx^2} + y = 0$.

(8) Let $F = R$, and let $V = R^n$ be the set of all real ordered n-tuples, $\alpha = (a_1, a_2, \ldots, a_n)$ with $a_i \in F$. Vector addition and scalar multiplication are defined by the rules

$$\begin{aligned}(a_1, \ldots, a_n) + (b_1, \ldots, b_n) &= (a_1 + b_1, \ldots, a_n + b_n),\\ a(a_1, \ldots, a_n) &= (aa_1, \ldots, aa_n).\end{aligned} \tag{1.2}$$

We call this vector space the *n-dimensional real coordinate space* or the *real affine n-space*. (The name "Euclidean n-space" is sometimes used, but that term should be reserved for an affine n-space in which distance is defined.)

(9) Let F^n be the set of all n-tuples of elements of F. Vector addition and scalar multiplication are defined by the rules (1.2). We call this vector space an *n-dimensional coordinate space*.

An immediate consequence of the axioms defining a vector space is that the zero vector, whose existence is asserted in $A3$, and the negative vector, whose existence is asserted in $A4$, are unique. Specifically, suppose 0 satisfies $A3$ for *all* vectors in V and that for *some* $\alpha \in V$ there is a $0'$ satisfying the condition $\alpha + 0 = \alpha + 0' = \alpha$. Then $0' = 0' + 0 = 0' + (\alpha + (-\alpha)) = (0' + \alpha) + (-\alpha) = (\alpha + 0') + (-\alpha) = \alpha + (-\alpha) = 0$. Notice that we have proved not merely that the zero vector satisfying $A3$ for *all* α is unique; we have proved that a vector satisfying the condition of $A3$ for *some* α must be the zero vector, which is a much stronger statement.

Also, suppose that to a given α there were two negatives, $(-\alpha)$ and $(-\alpha)'$, satisfying the conditions of $A4$. Then $(-\alpha)' = (-\alpha)' + 0 = (-\alpha)' + \alpha + (-\alpha) = (-\alpha) + \alpha + (-\alpha)' = (-\alpha) + 0 = (-\alpha)$. Both these demonstrations used the commutative law, $A5$. Use of this axiom could have been avoided, but the necessary argument would then have been somewhat longer.

End of Sentence

Uniqueness enables us to prove that $0\alpha = 0$. (Here is an example of the seemingly ambiguous use of the symbol "0." The "0" on the left side is a scalar while that on the right is a vector. However, no other interpretation could be given the symbols and it proves convenient to conform to the convention rather than introduce some other symbol for the zero vector.) For each $\alpha \in V$, $\alpha = 1 \cdot \alpha = (1 + 0)\alpha = 1 \cdot \alpha + 0 \cdot \alpha = \alpha + 0 \cdot \alpha$. Thus $0 \cdot \alpha = 0$. In a similar manner we can show that $(-1)\alpha = -\alpha]$ $\alpha + (-1)\alpha = (1 - 1)\alpha = 0 \cdot \alpha = 0$. Since the negative vector is unique we see that $(-1)\alpha = -\alpha$. It also follows similarly that $a \cdot 0 = 0$.

EXERCISES

1 to 4. What theorems must be proved in each of the Examples (4), (5), (6), and (7) to verify $A1$? To verify $B1$? (These axioms are usually the ones which require most specific verification. For example, if we establish that the vector space described in Example (3) satisfies all the axioms of a vector space, then $A1$ and $B1$ are the only ones that must be verified for Examples (4), (5), (6), and (7). Why?)

5. Let P^+ be the set of polynomials with real coefficients and positive constant term. Is P^+ a vector space? Why?

6. Show that if $a\alpha = 0$ and $a \neq 0$, then $\alpha = 0$. (*Hint:* Use axiom $F9$ for fields.)

7. Show that if $a\alpha = 0$ and $\alpha \neq 0$, then $a = 0$.

8. Show that the ξ such that $\alpha + \xi = \beta$ is (uniquely) $\xi = \beta + (-\alpha)$.

9. Let $\alpha = (2, -5, 0, 1)$ and $\beta = (-3, 3, 1, -1)$ be vectors in the coordinate space R^4. Determine

(*a*) $\alpha + \beta$.
(*b*) $\alpha - \beta$.
(*c*) 3α.
(*d*) $2\alpha + 3\beta$.

10. Show that any field can be considered to be a vector space over itself.

11. Show that the real numbers can be considered to be a vector space over the rational numbers.

12. Show that the complex numbers can be considered to be a vector space over the real numbers.

13. Prove the uniqueness of the zero vector and the uniqueness of the negative of each vector without using the commutative law, $A5$.

2 | Linear Independence and Linear Dependence

Because of the associative law for vector addition, we can omit the parentheses from expressions like $a_1\alpha_1 + (a_2\alpha_2 + a_3\alpha_3) = (a_1\alpha_1 + a_2\alpha_2) + a_3\alpha_3$ and write them in the simpler form $a_1\alpha_1 + a_2\alpha_2 + a_3\alpha_3 = \sum_{i=1}^{3} a_i\alpha_i$. It is clear that this convention can be extended to a sum of any number of such terms provided that only a finite number of coefficients are different from zero. Thus, whenever we write an expression like $\sum_i a_i\alpha_i$ (in which we do not specify the range of summation), it will be assumed, tacitly if not explicitly, that the expression contains only a finite number of non-zero coefficients.

If $\beta = \sum_i a_i\alpha_i$, we say that β is a *linear combination* of the α_i. We also say that β is *linearly dependent* on the α_i if β can be expressed as a linear combination of the α_i. An expression of the form $\sum_i a_i\alpha_i = 0$ is called a *linear relation* among the α_i. A relation with all $a_i = 0$ is called a *trivial linear relation*; a relation in which at least one coefficient is non-zero is called a *non-trivial linear relation.*

Definition. A set of vectors is said to be *linearly dependent* if there exists a non-trivial linear relation among them. Otherwise, the set is said to be *linearly independent.*

It should be noted that any set of vectors that includes the zero vector is linearly dependent. A set consisting of exactly one non-zero vector is linearly independent. For if $a\alpha = 0$ with $a \neq 0$, then $\alpha = 1 \cdot \alpha = (a^{-1} \cdot a)\alpha = a^{-1}(a\alpha) = a^{-1} \cdot 0 = 0$. Notice also that the empty set is linearly independent.

It is clear that the concept of linear independence of a set would be meaningless if a vector from a set could occur arbitrarily often in a possible relation. If a set of vectors is given, however, by itemizing the vectors in the set it is a definite inconvenience to insist that all the vectors listed be distinct. The burden of counting the number of times a vector can appear in a relation is transferred to the index set. For each index in the index set, we require that a linear relation contain but one term corresponding to that index. Similarly, when we specify a set by itemizing the vectors in the set, we require that one and only one vector be listed for each index in the index set. But we allow the possibility that several indices may be used to identify the same vector. Thus the set $\{\alpha_1, \alpha_2\}$, where $\alpha_1 = \alpha_2$ is linearly dependent, and any set with any vector listed at least twice is linearly dependent. To be precise, the concept of linear independence is a property of *indexed sets* and not a property of sets. In the example given above, the relation $\alpha_1 - \alpha_2 = 0$ involves two terms in the indexed set $\{\alpha_i \mid i \in \{1, 2\}\}$ while the set $\{\alpha_1, \alpha_2\}$ actually contains only one vector. We should refer to the linear dependence of an *indexed* set rather than the linear dependence of a set. The conventional terminology, which we are adopting, is inaccurate. This usage, however, is firmly rooted in tradition and, once understood, it is a convenience and not a source of difficulty. We speak of the linear dependence of a set, but the concept always refers to an indexed set. For a linearly independent indexed set, no vector can be listed twice; so in this case the inaccuracy of referring to a set rather than an indexed set is unimportant.

The concept of linear dependence and independence is used in essentially two ways. (1) If a set $\{\alpha_i\}$ of vectors is known to be linearly dependent, there exists a non-trivial linear relation of the form $\sum_i a_i\alpha_i = 0$. (This relation is not unique, but that is usually incidental.) There is at least one non-zero coefficient; let a_k be non-zero. Then $\alpha_k = \sum_{i\neq k} (-a_k^{-1}a_i)\alpha_i$; that is one of the vectors of the set $\{\alpha_i\}$ is a linear combination of the others. (2) If a set $\{\alpha_i\}$ of vectors is known to be linearly independent and a linear relation $\sum_i a_i\alpha_i = 0$ is obtained, we can conclude that all $a_i = 0$. This seemingly trivial observation is surprisingly useful.

In Example (1) the zero vector is the polynomial with all coefficients equal to zero. Thus the set of monomials $\{1, x, x^2, \ldots\}$ is a linearly independent set. The set $\{1, x, x^2, x^2 + x + 1\}$ is linearly dependent since $1 + x + x^2 - (x^2 + x + 1) = 0$. In P_n of Example (2), any $n + 1$ polynomials form a linearly dependent set.

In R^3 consider the vectors $\{\alpha = (1, 1, 0),\ \beta = (1, 0, 1),\ \gamma = (0, 1, 1),\ \delta = (1, 1, 1)\}$. These four vectors are linearly dependent since $\alpha + \beta + \gamma - 2\delta = 0$, yet any three of these four vectors are linearly independent.

Theorem 2.1. *If α is linearly dependent on $\{\beta_i\}$ and each β_i is linearly dependent on $\{\gamma_j\}$, then α is linearly dependent on $\{\gamma_j\}$.*

PROOF. From $\alpha = \sum_i b_i\beta_i$ and $\beta_i = \sum_j c_{ij}\gamma_j$ it follows that $\alpha = \sum_i b_i(\sum_j c_{ij}\gamma_j) = \sum_j (\sum_i b_ic_{ij})\gamma_j$. □

Theorem 2.2. *A set of non-zero vectors $\{\alpha_1, \alpha_2, \ldots\}$ is linearly dependent if and only if some α_k is a linear combination of the α_j with $j < k$.*

PROOF. Suppose the vectors $\{\alpha_1, \alpha_2, \ldots\}$ are linearly dependent. Then there is a non-trivial linear relation among them; $\sum_i a_i\alpha_i = 0$. Since a positive finite number of coefficients are non-zero, there is a last non-zero coefficient a_k. Furthermore, $k \geq 2$ since $\alpha_1 \neq 0$. Thus $\alpha_k = -a_k^{-1}\sum_{i=1}^{k-1} a_i\alpha_i = \sum_{i=1}^{k-1} (-a_k^{-1}\, a_i)\alpha_i$.

The converse is obvious. □

For any subset A of V the set of all linear combinations of vectors in A is called the set *spanned* by A, and we denote it by $\langle A\rangle$. We also say that A *spans* $\langle A\rangle$. It is a part of this definition that $A \subset \langle A\rangle$. We also agree that the empty set $\emptyset$ spans the set consisting of the zero vector alone. It is readily apparent that if $A \subset B$, then $\langle A\rangle \subset \langle B\rangle$.

In this notation Theorem 2.1 is equivalent to the statement: If $A \subset \langle B\rangle$ and $B \subset \langle C\rangle$, then $A \subset \langle C\rangle$.

Theorem 2.3. *The set $\{\alpha_i\}$ of non-zero vectors is linearly independent if and only if for each k, $\alpha_k \notin \langle\alpha_1, \ldots, \alpha_{k-1}\rangle$.* (To follow our definitions exactly, the set spanned by $\{\alpha_1, \ldots, \alpha_{k-1}\}$ should be denoted by $\langle\{\alpha_1, \ldots, \alpha_{k-1}\}\rangle$.

We shall use the symbol $\langle \alpha_1, \ldots, \alpha_{k-1} \rangle$ instead since it is simpler and there is no danger of ambiguity.)

PROOF. This is merely Theorem 2.2 in contrapositive form and stated in new notation. □

$\langle E \rangle \subset \langle F \rangle$

Theorem 2.4. *If* B *and* C *are any subsets such that* $B \subset \langle C \rangle$, *then* $\langle B \rangle \subset \langle C \rangle$.

PROOF. Set $A = \langle B \rangle$ in Theorem 2.1. Then $B \subset \langle C \rangle$ implies that $\langle B \rangle = A \subset \langle C \rangle$. □

Theorem 2.5. *If* $\alpha_k \in A$ *is dependent on the other vectors in* A, *then* $\langle A \rangle = \langle A - \{\alpha_k\} \rangle$.

page 22

PROOF. The assumption that α_k is dependent on $A - \{\alpha_k\}$ means that $A \subset \langle A - \{\alpha_k\} \rangle$. It then follows from Theorem 2.4 that $\langle A \rangle \subset \langle A - \{\alpha_k\} \rangle$. The equality follows from the fact that the inclusion in the other direction is evident. □

Theorem 2.6. *For any set* C, $\langle \langle C \rangle \rangle = \langle C \rangle$.

PROOF. Setting $B = \langle C \rangle$ in Theorem 2.4 we obtain $\langle \langle C \rangle \rangle = \langle B \rangle \subset \langle C \rangle$. Again, the inclusion in the other direction is obvious. □

Theorem 2.7. *If a finite set* $A = \{\alpha_1, \ldots, \alpha_n\}$ *spans* V, *then every linearly independent set contains at most n elements.*

PROOF. Let $B = \{\beta_1, \beta_2, \ldots\}$ be a linearly independent set. We shall successively replace the α_i by the β_j, obtaining at each step a new n-element set that spans V. Thus, suppose that $A_k = \{\beta_1, \ldots, \beta_k, \alpha_{k+1}, \ldots, \alpha_n\}$ is an n-element set that spans V. (Our starting point, the hypothesis of the theorem, is the case $k = 0$.) Since A_k spans V, β_{k+1} is dependent on A_k. Thus the set $\{\beta_1, \ldots, \beta_k, \beta_{k+1}, \alpha_{k+1}, \ldots, \alpha_n\}$ is linearly dependent. In any non-trivial relation that exists the non-zero coefficients cannot be confined to the β_j, for they are linearly independent. Thus one of the $\alpha_i (i > k)$ is dependent on the others, and after reindexing $\{\alpha_{k+1}, \ldots, \alpha_n\}$ if necessary we may assume that it is α_{k+1}. By Theorem 2.5 the set $A_{k+1} = \{\beta_1, \ldots, \beta_{k+1}, \alpha_{k+2}, \ldots, \alpha_n\}$ also spans V.

If there were more than n elements in B, we would in this manner arrive at the spanning set $A_n = \{\beta_1, \ldots, \beta_n\}$. But then the dependence of β_{n+1} on A_n would contradict the assumed linear independence of B. Thus B contains at most n elements. □

Theorem 2.7 is stated in slightly different forms in various books. The essential feature of the proof is the step-by-step replacement of the vectors in one set by the vectors in the other. The theorem is known as the Steinitz replacement theorem.

EXERCISES

1. In the vector space P of Example (1) let $p_1(x) = x^2 + x + 1$, $p_2(x) = x^2 - x - 2$, $p_3(x) = x^2 + x - 1$, $p_4(x) = x - 1$. Determine whether or not the set $\{p_1(x), p_2(x), p_3(x), p_4(x)\}$ is linearly independent. If the set is linearly dependent, express one element as a linear combination of the others.

2. Determine $\langle\{p_1(x), p_2(x), p_3(x), p_4(x)\}\rangle$, where the $p_i(x)$ are the same polynomials as those defined in Exercise 1. (The set required is infinite, so that we cannot list all its elements. What is required is a description; for example, "all polynomials of a certain degree or less," "all polynomials with certain kinds of coefficients," etc.)

3. A linearly independent set is said to be *maximal* if it is contained in no larger linearly independent set. In this definition the emphasis is on the concept of set inclusion and not on the number of elements in a set. In particular, the definition allows the possibility that two different maximal linearly independent sets might have different numbers of elements. Find all the maximal linearly independent subsets of the set given in Exercise 1. How many elements are in each of them?

4. Show that no finite set spans P; that is, show that there is no maximal finite linearly independent subset of P. Why are these two statements equivalent?

5. In Example (2) for $n = 4$, find a spanning set for P_4. Find a minimal spanning set. Use Theorem 2.7 to show that no other spanning set has fewer elements.

6. In Example (1) or (2) show that $\{1, x + 1, x^2 + x + 1, x^3 + x^2 + x + 1, x^4 + x^3 + x^2 + x + 1\}$ is a linearly independent set.

7. In Example (1) show that the set of all polynomials divisible by $x - 1$ cannot span P.

8. Determine which of the following set in R^4 are linearly independent over R.
(*a*) $\{(1, 1, 0, 1),\ (1, -1, 1, 1),\ (2, 2, 1, 2),\ (0, 1, 0, 0)\}$.
(*b*) $\{(1, 0, 0, 1),\ (0, 1, 1, 0),\ (1, 0, 1, 0),\ (0, 1, 0, 1)\}$.
(*c*) $\{(1, 0, 0, 1),\ (0, 1, 0, 1),\ (0, 0, 1, 1),\ (1, 1, 1, 1)\}$.

9. Show that $\{e_1 = (1, 0, 0, \dots, 0),\ e_2 = (0, 1, 0, \dots, 0), \dots,\ e_n = (0, 0, 0, \dots, 1)\}$ is linearly independent in F^n over F.

10. In Exercise 11 of Section 1 it was shown that we may consider the real numbers to be a vector space over the rational numbers. Show that $\{1, \sqrt{2}\}$ is a linearly independent set over the rationals. (This is equivalent to showing that $\sqrt{2}$ is irrational.) Using this result show that $\{1, \sqrt{2}, \sqrt{3}\}$ is linearly independent.

11. Show that if one vector of a set is the zero vector, then the set is linearly dependent.

12. Show that if an indexed set of vectors has one vector listed twice, the set is linearly dependent.

13. Show that if a subset of S is linearly dependent, then S is linearly dependent.

14. Show that if a set S is linearly independent, then every subset of S is linearly independent.

15. Show that if the set $A = \{\alpha_1, \ldots, \alpha_n\}$ is linearly independent and $\{\alpha_1, \ldots, \alpha_n, \beta\}$ is linearly dependent, then β is dependent on A.

16. Show that, if each of the vectors $\{\beta_0, \beta_1, \ldots, \beta_n\}$ is a linear combination of the vectors $\{\alpha_1, \ldots, \alpha_n\}$, then $\{\beta_0, \beta_1, \ldots, \beta_n\}$ is linearly dependent.

3 | Bases of Vector Spaces

Definition. A linearly independent set spanning a vector space V is called a *basis* or *base* (the plural is *bases*) of V.

If $A = \{\alpha_1, \alpha_2, \ldots\}$ is a basis of V, by definition an $\alpha \in V$ can be written in the form $\alpha = \sum_i a_i\alpha_i$. The interesting thing about a basis, as distinct from other spanning sets, is that the coefficients are uniquely determined by α. For suppose that we also have $\alpha = \sum_i b_i\alpha_i$. Upon subtraction we get the linear relation $\sum_i (a_i - b_i)\alpha_i = 0$. Since $\{\alpha_i\}$ is a linearly independent set, $a_i - b_i = 0$ and $a_i = b_i$ for each i. A related fact is that a basis is a particularly efficient spanning set, as we shall see.

In Example (1) the vectors $\{\alpha_i = x^i \mid i = 0, 1, \ldots\}$ form a basis. We have already observed that this set is linearly independent, and it clearly spans the space of all polynomials. The space P_n has a basis with a finite number of elements; $\{1, x, x^2, \ldots, x^{n-1}\}$.

The vector spaces in Examples (3), (4), (5), (6), and (7) do not have bases with a finite number of elements.

In Example (8) every R^n has a finite basis consisting of $\{\alpha_i \mid \alpha_i = (\delta_{1i}, \delta_{2i}, \ldots, \delta_{ni})\}$. (Here δ_{ij} is the useful symbol known as the Kronecker delta. By definition $\delta_{ij} = 0$ if $i \neq j$ and $\delta_{ii} = 1$.)

Theorem 3.1. *If a vector space has one basis with a finite number of elements, then all other bases are finite and have the same number of elements.*

PROOF. Let A be a basis with a finite number n of elements, and let B be any other basis. Since A spans V and B is linearly independent, by Theorem 2.7 the number m of element in B must be at most n. This shows that B is finite and $m \leq n$. But then the roles of A and B can be interchanged to obtain the inequality in the other order so that $m = n$. □

A vector space with a finite basis is called a *finite dimensional* vector space, and the number of elements in a basis is called the *dimension* of the space. Theorem 3.1 says that the dimension of a finite dimensional vector space is well defined. The vector space with just one element, the zero vector, has one linearly independent subset, the empty set. The empty set is also a spanning set and is therefore a basis of $\{0\}$. Thus $\{0\}$ has dimension zero. There are very interesting vector spaces with infinite bases; for example, P of Example (1). Moreover, many of the theorems and proofs we give are also valid for infinite dimensional vector spaces. It is not our intention, however, to deal with infinite dimensional vector spaces as such, and whenever we speak of the dimension of a vector space without specifying whether it is finite or infinite dimensional we mean that the dimension is finite.

Among the examples we have discussed so far, each P_n and each R^n is n-dimensional. We have already given at least one basis for each. There are many others. The bases we have given happen to be conventional and convenient choices.

Theorem 3.2. *Any $n + 1$ vectors in an n-dimensional vector space are linearly dependent.*

PROOF. Their independence would contradict Theorem 2.7. □

We have already seen that the four vectors $\{\alpha = (1, 1, 0), \beta = (1, 0, 1), \gamma = (0, 1, 1), \delta = (1, 1, 1)\}$ form a linearly dependent set in R^3. Since R^3 is 3-dimensional we see that this must be expected for any set containing at least four vectors from R^3. The next theorem shows that each subset of three is a basis.

Theorem 3.3. *A set of n vectors in an n-dimensional vector space V is a basis if and only if it is linearly independent.* i.e. a set of n linearly independent vectors must span V.

PROOF. The "only if" is part of the definition of a basis. Let $A = \{\alpha_1, \ldots, \alpha_n\}$ be a linearly independent set and let α be any vector in V. Since $\{\alpha_1, \ldots, \alpha_n, \alpha\}$ contains $n + 1$ elements it must be linearly dependent. Any non-trivial relation that exists must contain α with a non-zero coefficient, for if that coefficient were zero the relation would amount to a relation in A. Thus α is dependent on A. Hence A spans V and is a basis. □

Theorem 3.4. *A set of n vectors in an n-dimensional vector space V is a basis if and only if it spans V.* i.e. a set of n vectors spanning V must be linearly independent

PROOF. The "only if" is part of the definition of a basis. If n vectors did span V and were linearly dependent, then (by Theorem 2.5) a proper subset would also span V, ~~contrary to Theorem 2.7. □~~ * See below

We see that a basis is a maximal linearly independent set and a minimal spanning set. This idea is made explicit in the next two theorems.

Theorem 3.5. *In a finite dimensional vector space, every spanning set contains a basis.*

PROOF. Let B be a set spanning V. If $V = \{0\}$, then $\emptyset \subset B$ is a basis of $\{0\}$. If $V \neq \{0\}$, then B must contain at least one non-zero vector α_1. We now search for another vector in B which is not dependent on $\{\alpha_1\}$. We call this vector α_2 and search for another vector in B which is not dependent on the linearly independent set $\{\alpha_1, \alpha_2\}$. We continue in this way as long as we can, but the process must terminate as we cannot find more than n

* In any n dimensional space any spanning set must contain at least n vector See unit 1.3.2. If the spanning set contain more than n vectors then by Theorem 3.2 above we can find linearly independent vectors.

Because the space is n [illegible] [illegible] which
spans the space has n vectors [illegible] Theorem 2.7

linearly independent vectors in B. Thus suppose we have obtained the set $A = \{\alpha_1, \ldots, \alpha_m\}$ with the property that every vector in B is linearly dependent on A. Then because of Theorem 2.1 the set A must also span V and it is a basis. □

To drop the assumption that the vector space is n-dimensional would change the complexion of Theorem 3.5 entirely. As it stands the theorem is interesting but minor, and not difficult to prove. Without this assumption the theorem would assert that *every* vector space has a basis since every vector space is spanned by itself. Discussion of such a theorem is beyond the aims of this treatment of the subject of vector spaces.

Theorem 3.6. *In a finite dimensional vector space any linearly independent set of vectors can be extended to a basis.*

PROOF. Let $A = \{\alpha_1, \ldots, \alpha_n\}$ be a basis of V, and let $B = \{\beta_1, \ldots, \beta_m\}$ be a linearly independent set ($m \leq n$). The set $\{\beta_1, \ldots, \beta_m, \alpha_1, \ldots, \alpha_n\}$ spans V. If this set is linearly dependent (and it surely is if $m > 0$) then some element is a linear combination of the preceding elements (Theorem 2.2). This element cannot be one of the β_i's for then B would be linearly dependent. But then this α_i can be removed to obtain a smaller set spanning V (Theorem 2.5). We continue in this way, discarding elements as long as we have a linearly dependent spanning set. At no stage do we discard one of the β_i's. Since our spanning set is finite this process must terminate with a basis containing B as a subset. □

Theorem 3.6 is one of the most frequently used theorems in the book. It is often used in the following way. A non-zero vector with a certain desired property is selected. Since the vector is non-zero, the set consisting of that vector alone is a linearly independent set. An application of Theorem 3.6 shows that there is a basis containing that vector. This is usually the first step of a proof by induction in which a basis is obtained for which all the vectors in the basis have the desired property.

Let $A = \{\alpha_1, \ldots, \alpha_n\}$ be an arbitrary basis of V, a vector space of dimension n over the field F. Let α be any vector in V. Since A is a spanning set α can be represented as a linear combination of the form $\alpha = \sum_{i=1}^n a_i\alpha_i$. Since A is linearly independent this representation is unique, that is, the coefficients a_i are uniquely determined by α (for the given basis A). On the other hand, for each n-tuple $(a_1, \ldots, a_n)$ there is a vector in V of the form $\sum_{i=1}^n a_i\alpha_i$. Thus there is a one-to-one correspondence between the vectors in V and the n-tuples $(a_1, \ldots, a_n) \in F^n$.

If $\alpha = \sum_{i=1}^n a_i\alpha_i$, the scalar a_i is called the *i-th coordinate* of α, and $a_i\alpha_i$ is called the *i-th component* of α. Generally, coordinates and components depend on the choice of the entire basis and cannot be determined from individual vectors in the basis. Because of the rather simple correspondence between coordinates and components there is a tendency to confuse them and to use both terms for both concepts. Since the intended meaning is usually clear from context, this is seldom a source of difficulty.

If $\alpha = \sum_{i=1}^{n} a_i\alpha_i$ corresponds to the n-tuple $(a_1, \ldots, a_n)$ and $\beta = \sum_{i=1}^{n} b_i\alpha_i$ corresponds to the n-tuple $(b_1, \ldots, b_n)$, then $\alpha + \beta = \sum_{i=1}^{n} (a_i + b_i)\alpha_i$ corresponds to the n-tuple $(a_1 + b_1, \ldots, a_n + b_n)$. Also, $a\alpha = \sum_{i=1}^{n} aa_i\alpha_i$ corresponds to the n-tuple $(aa_1, \ldots, aa_n)$. Thus the definitions of vector addition and scalar multiplication among n-tuples defined in Example (9) correspond exactly to the corresponding operations in V among the vectors which they represent. When two sets of objects can be put into a one-to-one correspondence which preserves all significant relations among their elements, we say the two sets are *isomorphic*; that is, they have the same form. Using this terminology, we can say that every vector space of dimension n over a given field F is isomorphic to the n-dimensional coordinate space F^n. Two sets which are isomorphic differ in details which are not related to their internal structure. They are essentially the same. Furthermore, since two sets isomorphic to a third are isomorphic to each other we see that all n-dimensional vector spaces over the same field of scalars are isomorphic.

The set of n-tuples together with the rules for addition and scalar multiplication forms a vector space in its own right. However, when a basis is chosen in an abstract vector space V the correspondence described above establishes an isomorphism between V and F^n. In this context we consider F^n to be a *representation* of V. Because of the existence of this isomorphism a study of vector spaces could be confined to a study of coordinate spaces. However, the exact nature of the correspondence between V and F^n depends upon the choice of a basis in V. If another basis were chosen in V a correspondence between the $\alpha \in V$ and the n-tuples would exist as before, but the correspondence would be quite different. We choose to regard the vector space V and the vectors in V as the basic concepts and their representation by n-tuples as a tool for computation and convenience. There are two important benefits from this viewpoint. Since we are free to choose the basis we can try to choose a coordinatization for which the computations are particularly simple or for which some fact that we wish to demonstrate is particularly evident. In fact, the choice of a basis and the consequences of a change in basis is the central theme of matrix theory. In addition, this distinction between a vector and its representation removes the confusion that always occurs when we define a vector as an n-tuple and then use another n-tuple to represent it.

Only the most elementary types of calculations can be carried out in the abstract. Elaborate or complicated calculations usually require the introduction of a representing coordinate space. In particular, this will be required extensively in the exercises in this text. But the introduction of coordinates can result in confusions that are difficult to clarify without extensive verbal description or awkward notation. Since we wish to avoid cumbersome notation and keep descriptive material at a minimum in the exercises, it is helpful to spend some time clarifying conventional notations and circumlocutions that will appear in the exercises.

The introduction of a coordinate representation for V involves the selection of a basis $\{\alpha_1, \dots, \alpha_n\}$ for V. With this choice α_1 is represented by $(1, 0, \dots, 0)$, α_2 is represented by $(0, 1, 0, \dots, 0)$, etc. While it may be necessary to find a basis with certain desired properties the basis that is introduced at first is arbitrary and serves only to express whatever problem we face in a form suitable for computation. Accordingly, it is customary to suppress specific reference to the basis given initially. In this context it is customary to speak of "the vector $(a_1, a_2, \dots, a_n)$" rather than "the vector α whose representation with respect to the given basis $\{\alpha_1, \dots, \alpha_n\}$ is $(a_1, a_2, \dots, a_n)$." Such short-cuts may be disgracefully inexact, but they are so common that we must learn how to interpret them.

For example, let V be a two-dimensional vector space over R. Let $A = \{\alpha_1, \alpha_2\}$ be the selected basis. If $\beta_1 = \alpha_1 + \alpha_2$ and $\beta_2 = -\alpha_1 + \alpha_2$, then $B = \{\beta_1, \beta_2\}$ is also a basis of V. With the convention discussed above we would identify α_1 with $(1, 0)$, α_2 with $(0, 1)$, β_1 with $(1, 1)$, and β_2 with $(-1, 1)$. Thus, we would refer to the basis $B = \{(1, 1), (-1, 1)\}$. Since $\alpha_1 = \frac{1}{2}\beta_1 - \frac{1}{2}\beta_2$, α_1 has the representation $(\frac{1}{2}, -\frac{1}{2})$ with respect to the basis B. If we are not careful we can end up by saying that "$(1, 0)$ is represented by $(\frac{1}{2}, -\frac{1}{2})$."

EXERCISES

To show that a given set is a basis by direct appeal to the definition means that we must show the set is linearly independent and that it spans V. In any given situation, however, the task is very much simpler. Since V is n-dimensional a proposed basis must have n elements. Whether this is the case can be told at a glance. In view of Theorems 3.3 and 3.4 if a set has n elements, to show that it is a basis it suffices to show either that it spans V or that it is linearly independent.

1. In R^3 show that $\{(1, 1, 0), (1, 0, 1), (0, 1, 1)\}$ is a basis by showing that it is linearly independent.

2. Show that $\{(1, 1, 0), (1, 0, 1), (0, 1, 1)\}$ is a basis by showing that $\langle(1, 1, 0), (1, 0, 1), (0, 1, 1)\rangle$ contains $(1, 0, 0)$, $(0, 1, 0)$ and $(0, 0, 1)$. Why does this suffice?

3. In R^4 let $A = \{(1, 1, 0, 0), (0, 0, 1, 1), (1, 0, 1, 0), (0, 1, 0, -1)\}$ be a basis (is it?) and let $B = \{(1, 2, -1, 1), (0, 1, 2, -1)\}$ be a linearly independent set (is it?). Extend B to a basis of R^4. (There are many ways to extend B to a basis. It is intended here that the student carry out the steps of the proof of Theorem 3.6 for this particular case.)

4. Find a basis of R^4 containing the vector (1, 2, 3, 4). (This is another even simpler application of the proof of Theorem 3.6. This, however, is one of the most important applications of this theorem, to find a basis containing a particular vector.)

5. Show that a maximal linearly independent set is a basis.

6. Show that a minimal spanning set is a basis.

4 | Subspaces

Definition. A *subspace* W of a vector space V is a non-empty subset of V which is itself a vector space with respect to the operations of addition and scalar multiplication defined in V. In particular, the subspace must be a vector space over the same field F.

The first problem that must be settled is the problem of determining the conditions under which a subset W is in fact a subspace. It should be clear that axioms $A2$, $A5$, $B2$, $B3$, $B4$, and $B5$ need not be checked as they are valid in any subset of V. The most innocuous conditions seem to be $A1$ and $B1$, but it is precisely these conditions that must be checked. If $B1$ holds for a non-empty subset W, there is an $\alpha \in W$ so that $0\alpha = 0 \in W$. Also, for each $\alpha \in W$, $(-1)\alpha = -\alpha \in W$. Thus $A3$ and $A4$ follow from $B1$ in any non-empty subset of a vector space and it is sufficient to check that W is non-empty and closed under addition and scalar multiplication.

The two closure conditions can be combined into one statement: if α, $\beta \in W$ and a, $b \in F$, then $a\alpha + b\beta \in W$. This may seem to be a small change, but it is a very convenient form of the conditions. It is also equivalent to the statement that all linear combinations of elements in W are also in W; that is, $\langle W \rangle = W$. It follows directly from this statement that for any subset A, $\langle A \rangle$ is a subspace. Thus, instead of speaking of the subset spanned by A, we speak of the subspace spanned by A.

Every vector space V has V and the zero space $\{0\}$ as subspaces. As a rule we are interested in subspaces other than these and to distinguish them we call the subspaces other than V and $\{0\}$ *proper* subspaces. In addition, if W is a subspace we designate subspaces of W other than W and $\{0\}$ as proper subspaces of W.

In Examples (1) and (2) we can take a fixed finite set $\{x_1, x_2, \ldots, x_m\}$ of elements of F and define W to be the set of all polynomials such that $p(x_1) = p(x_2) = \cdots = p(x_m) = 0$. To show that W is a subspace it is sufficient to show that the sum of two polynomials which vanish at the x_i also vanishes at the x_i, and the product of a scalar and a polynomial vanishing at the x_i also vanishes at the x_i. What is the situation in P_n if $m > n$? Similar subspaces can be defined in examples (3), (4), (5), (6), and (7).

The space P_m is a subspace of P, and also a subspace of P_n for $m \leq n$.

In R^n, for each m, $0 \leq m \leq n$, the set of all $\alpha = (a_1, a_2, \dots, a_n)$ such that $a_1 = a_2 = \cdots = a_m = 0$ is a subspace of R^n. This subspace is proper if $0 < m < n$.

Notice that the set of all n-tuples of rational numbers is a subset of R^n and it is a vector space over the rational numbers, but it is not a subspace of R^n since it is not a vector space over the real numbers. Why?

Theorem 4.1. *The intersection of any collection of subspaces is a subspace.*

PROOF. Let $W_\mu : \mu \in M$ be an indexed collection of subspaces of V. $\cap_{\mu \in M} W_\mu$ is not empty since it contains 0. Let α, $\beta \in \cap_{\mu \in M} W_\mu$ and $a, b \in F$. Then α, $\beta \in W_\mu$ for each $\mu \in M$. Since W_μ is a subspace $a\alpha + b\beta \in W_\mu$ for each $\mu \in M$, and hence $a\alpha + b\beta \in \cap_{\mu \in M} W_\mu$. Thus $\cap_{\mu \in M} W_\mu$ is a subspace. □

Let A be any subset of V, not necessarily a subspace. There exist subspaces $W_\mu \subset V$ which contain A; in fact, V is one of them. The intersection $\cap_{A \subset W_\mu} W_\mu$ of all such subspaces is a subspace containing A. It is the smallest subspace containing A.

Theorem 4.2. *For any $A \subset V$, $\cap_{A \subset W_\mu} W_\mu = \langle A \rangle$; that is, the smallest subspace containing A is exactly the subspace spanned by A.*

PROOF. Since $\cap_{A \subset W_\mu} W_\mu$ is a subspace containing A, it contains all linear combinations of elements of A. Thus $\langle A \rangle \subset \cap_{A \subset W_\mu} W_\mu$. On the other hand $\langle A \rangle$ is a subspace containing A, that is, $\langle A \rangle$ is one of the W_μ and hence $\cap_{A \subset W_\mu} W_\mu \subset \langle A \rangle$. Thus $\cap_{A \subset W_\mu} W_\mu = \langle A \rangle$. □

$W_1 + W_2$ is defined to be the set of all vectors of the form $\alpha_1 + \alpha_2$ where $\alpha_1 \in W_1$ and $\alpha_2 \in W_2$.

Theorem 4.3. *If W_1 and W_2 are subspaces of V, then $W_1 + W_2$ is a subspace of V.*

PROOF. If $\alpha = \alpha_1 + \alpha_2 \in W_1 + W_2$, $\beta = \beta_1 + \beta_2 \in W_1 + W_2$, and a, $b \in F$, then $a\alpha + b\beta = a(\alpha_1 + \alpha_2) + b(\beta_1 + \beta_2) = (a\alpha_1 + b\beta_1) + (a\alpha_2 + b\beta_2) \in W_1 + W_2$. Thus $W_1 + W_2$ is a subspace. □

Theorem 4.4. *$W_1 + W_2$ is the smallest subspace containing both W_1 and W_2; that is, $W_1 + W_2 = \langle W_1 \cup W_2 \rangle$. If A_1 spans W_1 and A_2 spans W_2, then $A_1 \cup A_2$ spans $W_1 + W_2$.*

PROOF. Since $0 \in W_1$, $W_2 \subset W_1 + W_2$. Similarly, $W_1 \subset W_1 + W_2$. Since $W_1 + W_2$ is a subspace containing $W_1 \cup W_2$, $\langle W_1 \cup W_2 \rangle \subset W_1 + W_2$. For any $\alpha \in W_1 + W_2$, α can be written in the form $\alpha = \alpha_1 + \alpha_2$ where $\alpha_1 \in W_1$ and $\alpha_2 \in W_2$. Then $\alpha_1 \in W_1 \subset \langle W_1 \cup W_2 \rangle$ and $\alpha_2 \in W_2 \subset \langle W_1 \cup W_2 \rangle$. Since $\langle W_1 \cup W_2 \rangle$ is a subspace, $\alpha = \alpha_1 + \alpha_2 \in \langle W_1 \cup W_2 \rangle$. Thus $W_1 + W_2 = \langle W_1 \cup W_2 \rangle$.

The second part of the theorem now follows directly. $W_1 = \langle A_1 \rangle \subset \langle A_1 \cup A_2 \rangle$ and $W_2 = \langle A_2 \rangle \subset \langle A_1 \cup A_2 \rangle$ so that $W_1 \cup W_2 \subset \langle A_1 \cup A_2 \rangle \subset \langle W_1 \cup W_2 \rangle$, and hence $\langle W_1 \cup W_2 \rangle = \langle A_1 \cup A_2 \rangle$. □

Theorem 4.5. *A subspace W of an n-dimensional vector space V is a finite dimensional vector space of dimension $m \leq n$.*

PROOF. If $W = \{0\}$, then W is 0-dimensional. Otherwise, there is a non-zero vector $\alpha_1 \in W$. If $\langle \alpha_1 \rangle = W$, W is 1-dimensional. Otherwise, there is an $\alpha_2 \notin \langle \alpha_1 \rangle$ in W. We continue in this fashion as long as possible. Suppose we have obtained the linearly independent set $\{\alpha_1, \ldots, \alpha_k\}$ and that it does not span W. Then there exists an $\alpha_{k+1} \in W$, $\alpha_{k+1} \notin \langle \alpha_1, \ldots, \alpha_k \rangle$. In a linear relation of the form $\sum_{i=1}^{k+1} a_i\alpha_i = 0$ we could not have $a_{k+1} \neq 0$ for then $\alpha_{k+1} \in \langle \alpha_1, \ldots, \alpha_k \rangle$. But then the relation reduces to the form $\sum_{i=1}^{k} a_i\alpha_i = 0$. Since $\{\alpha_1, \ldots, \alpha_k\}$ is linearly independent, all $a_i = 0$. Thus $\{\alpha_1, \ldots, \alpha_k, \alpha_{k+1}\}$ is linearly independent. In general, any linearly independent set in W that does not span W can be expanded into a larger linearly independent set in W. This process cannot go on indefinitely for in that event we would obtain more than n linearly independent vectors in V. Thus there exists an m such that $\langle \alpha_1, \ldots, \alpha_m \rangle = W$. It is clear that $m \leq n$. □

Theorem 4.6. *Given any subspace W of dimension m in an n-dimensional vector space V, there exists a basis $\{\alpha_1, \ldots, \alpha_m, \alpha_{m+1}, \ldots, \alpha_n\}$* of V *such that $\{\alpha_1, \ldots, \alpha_m\}$ is a basis of W.*

PROOF. By the previous theorem we see that W has a basis $\{\alpha_1, \ldots, \alpha_m\}$. This set is also linearly independent when considered in V, and hence by Theorem 3.6 it can be extended to a basis of V. □

Theorem 4.7. *If two subspaces U and W of a vector space V have the same finite dimension and $U \subset W$, then $U = W$.*

PROOF. By the previous theorem there exists a basis of U which can be extended to a basis of W. But since $\dim U = \dim W$, the basis of W can have no more elements than does the basis of U. This means a basis of U is also a basis of W; that is, $U = W$. □

Theorem 4.8. *If W_1 and W_2 are any two subspaces of a finite dimensional vector space V, then $\dim (W_1 + W_2) = \dim W_1 + \dim W_2 - \dim (W_1 \cap W_2)$.*

PROOF. Let $\{\alpha_1, \ldots, \alpha_r\}$ be a basis of $W_1 \cap W_2$. This basis can be extended to a basis $\{\alpha_1, \ldots, \alpha_r, \beta_1, \ldots, \beta_s\}$ of W_1 and also to a basis $\{\alpha_1, \ldots, \alpha_r, \gamma_1, \ldots, \gamma_t\}$ of W_2. It is clear that $\{\alpha_1, \ldots, \alpha_r, \beta_1, \ldots, \beta_s, \gamma_1, \ldots, \gamma_t\}$ spans $W_1 + W_2$; we wish to show that this set is linearly independent. Suppose $\sum_i a_i\alpha_i + \sum_j b_j\beta_j + \sum_k c_k\gamma_k = 0$ is a linear relation. Then $\sum_i a_i\alpha_i + \sum_j b_j\beta_j = -\sum_k c_k\gamma_k$. The left side is in W_1 and the right side is in W_2, and hence both are in $W_1 \cap W_2$. Each side is then expressible as a linear combination of the $\{\alpha_i\}$. Since any representation of an element as a linear combination of the $\{\alpha_1, \ldots, \alpha_r, \beta_1, \ldots, \beta_s\}$ is unique, this means that

$b_j = 0$ for all j. By a symmetric argument we see that all $c_k = 0$. Finally, this means that $\sum_i a_i\alpha_i = 0$ from which it follows that all $a_i = 0$. This shows that the spanning set $\{\alpha_1, \ldots, \alpha_r, \beta_1, \ldots, \beta_s, \gamma_1, \ldots, \gamma_t\}$ is linearly independent and a basis of $W_1 + W_2$. Thus $\dim (W_1 + W_2) = r + s + t = (r + s) + (r + t) - r = \dim W_1 + \dim W_2 - \dim (W_1 \cap W_2)$. □

As an example, consider in R^3 the subspaces $W_1 = \langle(1, 0, 2), (1, 2, 2)\rangle$ and $W_2 = \langle(1, 1, 0), (0, 1, 1)\rangle$. Both subspaces are of dimension 2. Since $W_1 \subset W_1 + W_2 \subset R^3$ we see that $2 \leq \dim (W_1 + W_2) \leq 3$. Because of Theorem 4.8 this implies that $1 \leq \dim (W_1 \cap W_2) \leq 2$. In more familiar terms, W_1 and W_2 are planes in a 3-dimensional space. Since both planes contain the origin, they do intersect. Their intersection is either a line or, in case they coincide, a plane. The first problem is to find a basis for $W_1 \cap W_2$. Any $\alpha \in W_1 \cap W_2$ must be expressible in the forms $\alpha = a(1, 0, 2) + b(1, 2, 2) = c(1, 1, 0) + d(0, 1, 1)$. This leads to the three equations:

$$\begin{aligned} a + b &= c \\ 2b &= c + d \\ 2a + 2b &= \phantom{c + {}} d. \end{aligned}$$

These equations have the solutions $b = -3a$, $c = -2a$, $d = -4a$. Thus $\alpha = a(1, 0, 2) - 3a(1, 2, 2) = a(-2, -6, -4)$. As a check we also have $\alpha = -2a(1, 1, 0) - 4a(0, 1, 1) = a(-2, -6, -4)$. We have determined that $\{(1, 3, 2)\}$ is a basis of $W_1 \cap W_2$. Also $\{(1, 3, 2), (1, 0, 2)\}$ is a basis of W_1 and $\{(1, 3, 2), (1, 1, 0)\}$ is a basis of W_2.

We are all familiar with the theorem from solid geometry to the effect that two non-parallel planes intersect in a line, and the example above is an illustration of that theorem. In spaces of dimension higher than 3, however, it is possible for two subspaces of dimension 2 to have but one point in common. For example, in R^4 the subspaces $W_1 = \langle(1, 0, 0, 0), (0, 1, 0, 0)\rangle$ and $W_2 = \langle(0, 0, 1, 0), (0, 0, 0, 1)\rangle$ are each 2-dimensional and $W_1 \cap W_2 = \{0\}$, $W_1 + W_2 = R^4$.

Those cases in which $\dim (W_1 \cap W_2) = 0$ deserve special mention. If $W_1 \cap W_2 = \{0\}$ we say that the sum $W_1 + W_2$ is *direct*: $W_1 + W_2$ is a *direct sum* of W_1 and W_2. To indicate that a sum is direct we use the notation, $W_1 \oplus W_2$. For $\alpha \in W_1 \oplus W_2$ there exist $\alpha_1 \in W_1$ and $\alpha_2 \in W_2$ such that $\alpha = \alpha_1 + \alpha_2$. This much is true for any sum of two subspaces. If the sum is direct, however, α_1 and α_2 are uniquely determined by α. For if $\alpha = \alpha_1 + \alpha_2 = \alpha_1' + \alpha_2'$, then $\alpha_1 - \alpha_1' = \alpha_2' - \alpha_2$. Since the left side is in W_1 and the right side is in W_2, both are in $W_1 \cap W_2$. But this means $\alpha_1 - \alpha_1' = 0$ and $\alpha_2 - \alpha_2 = 0$; that is, the decomposition of α into a sum of an element in W_1 plus an element in W_2 is unique. If V is the direct sum of W_1 and W_2, we say that W_1 and W_2 are *complementary* and that W_2 is a *complementary subspace* of W_1, or a *complement* of W_1.

The notion of a direct sum can be extended to a sum of any finite number of subspaces. The sum $W_1 + \cdots + W_k$ is said to be *direct* if for each i, $W_i \cap (\sum_{j \neq i} W_j) = \{0\}$. If the sum of several subspaces is direct, we use the notation $W_1 \oplus W_2 \oplus \cdots \oplus W_k$. In this case, too, $\alpha \in W_1 \oplus \cdots \oplus W_k$ can be expressed uniquely in the form $\alpha = \sum_i \alpha_i$, $\alpha_i \in W_i$.

Theorem 4.9. *If W is a subspace of V there exists a subspace W' such that $V = W \oplus W'$.*

PROOF. Let $\{\alpha_1, \ldots, \alpha_m\}$ be a basis of W. Extend this linearly independent set to a basis $\{\alpha_1, \ldots, \alpha_m, \alpha_{m+1}, \ldots, \alpha_n\}$ of V. Let W' be the subspace spanned by $\{\alpha_{m+1}, \ldots, \alpha_n\}$. Clearly, $W \cap W' = \{0\}$ and the sum $V = W + W'$ is direct. □

Thus every subspace of a finite dimensional vector space has a complementary subspace. The complement is not unique, however. If for W there exists a subspace W' such that $V = W \oplus W'$, we say that W is a *direct summand* of V.

Theorem 4.10. *For a sum of several subspaces of a finite dimensional vector space to be direct it is necessary and sufficient that* $\dim(W_1 + \cdots + W_k) = \dim W_1 + \cdots + \dim W_k$.

PROOF. This is an immediate consequence of Theorem 4.8 and the principle of mathematical induction. □

EXERCISES

1. Let P be the space of all polynomials with real coefficients. Determine which of the following subsets of P are subspaces.

(*a*) $\{p(x) \mid p(1) = 0\}$.
(*b*) $\{p(x) \mid \text{constant term of } p(x) = 0\}$.
(*c*) $\{p(x) \mid \text{degree of } p(x) = 3\}$.
(*d*) $\{p(x) \mid \text{degree of } p(x) \leq 3\}$.

(Strictly speaking, the zero polynomial does not have a degree associated with it. It is sometimes convenient to agree that the zero polynomial has degree less than any integer, positive or negative. With this convention the zero polynomial is included in the set described above, and it is not necessary to add a separate comment to include it.)

(*e*) $\{p(x) \mid \text{degree of } p(x) \text{ is even}\} \cup \{0\}$.

2. Determine which of the following subsets of R^n are subspaces.

(*a*) $\{(x_1, x_2, \ldots, x_n) \mid x_1 = 0\}$.
(*b*) $\{(x_1, x_2, \ldots, x_n) \mid x_1 \geq 0\}$.
(*c*) $\{(x_1, x_2, \ldots, x_n) \mid x_1 + 2x_2 = 0\}$.
(*d*) $\{(x_1, x_2, \ldots, x_n) \mid x_1 + 2x_2 = 1\}$.
(*e*) $\{(x_1, x_2, \ldots, x_n) \mid x_1 + 2x_2 \geq 0\}$.
(*f*) $\{(x_1, x_2, \ldots, x_n) \mid m_i < x_i < M_i$: $i = 1, 2, \ldots, n$ where the m_i and M_i are constants$\}$.
(*g*) $\{(x_1, x_2, \ldots, x_n) \mid x_1 = x_2 = \cdots = x_n\}$.

3. What is the essential difference between the condition used to define the subset in (*c*) of Exercise 2 and the condition used in (*d*)? Is the lack of a non-zero constant term important in (*c*)?

4. What is the essential difference between the condition used to define the subset in (*c*) of Exercise 2 and the condition used in (*e*)? What, in general, are the differences between the conditions in (*a*), (*c*), and (*g*) and those in (*b*), (*e*), and (*f*)?

5. Show that $\{(1, 1, 0, 0), (1, 0, 1, 1)\}$ and $\{(2, -1, 3, 3), (0, 1, -1, -1)\}$ span the same subspace of R^4.

6. Let W be the subspace of R^5 spanned by $\{(1, 1, 1, 1, 1), (1, 0, 1, 0, 1), (0, 1, 1, 1, 0), (2, 0, 0, 1, 1), (2, 1, 1, 2, 1), (1, -1, -1, -2, 2), (1, 2, 3, 4, -1)\}$. Find a basis for W and the dimension of W.

7. Show that $\{(1, -1, 2, -3), (1, 1, 2, 0), (3, -1, 6, -6)\}$ and $\{(1, 0, 1, 0), (0, 2, 0, 3)\}$ do not span the same subspace.

8. Let $W = \langle(1, 2, 3, 6), (4, -1, 3, 6), (5, 1, 6, 12)\rangle$ and $W_2 = \langle(1, -1, 1, 1), (2, -1, 4, 5)\rangle$ be subspaces of R^4. Find bases for $W_1 \cap W_2$ and $W_1 + W_2$. Extend the basis of $W_1 \cap W_2$ to a basis of W_1, and extend the basis of $W_1 \cap W_2$ to a basis of W_2. From these bases obtain a basis of $W_1 + W_2$.

9. Let P be the space of all polynomials with real coefficients, and let $W_1 = \{p(x) \mid p(1) = 0\}$ and $W_2 = \{p(x) \mid p(2) = 0\}$. Determine $W_1 \cap W_2$ and $W_1 + W_2$. (These spaces are infinite dimensional and the student is not expected to find bases for these subspaces. What is expected is a simple criterion or description of these subspaces.)

10. We have already seen (Section 1, Exercise 11) that the real numbers form a vector space over the rationals. Show that $\{1, \sqrt{2}\}$ and $\{1 - \sqrt{2}, 1 + \sqrt{2}\}$ span the same subspace.

11. Show that if W_1 and W_2 are subspaces, then $W_1 \cup W_2$ is not a subspace unless one is a subspace of the other.

12. Show that the set of all vectors $(x_1, x_2, x_3, x_4) \in R^4$ satisfying the equations

$$\begin{aligned} 3x_1 - 2x_2 - x_3 - 4x_4 &= 0 \\ x_1 + x_2 - 2x_3 - 3x_4 &= 0 \end{aligned}$$

is a subspace of R^4. Find a basis for this subspace. (*Hint:* Solve the equations for x_1 and x_2 in terms of x_3 and x_4. Then specify various values for x_3 and x_4 to obtain as many linearly independent vectors as are needed.)

13. Let S, T, and T^* be three subspaces of V (of finite dimension) for which (*a*) $S \cap T = S \cap T^*$, (*b*) $S + T = S + T^*$, (*c*) $T \subset T^*$. Show that $T = T^*$.

14. Show by example that it is possible to have $S \oplus T = S \oplus T^*$ without having $T = T^*$.

15. If $V = W_1 \oplus W_2$ and W is any subspace of V such that $W_1 \subset W$, show that $W = (W \cap W_1) \oplus (W \cap W_2)$. Show by an example that the condition $W_1 \subset W$ (or $W_2 \subset W$) is necessary.