Mathematics Course 111: Algebra I
Part IV: Vector Spaces

D. R. Wilkins

Academic Year 1996-7

9 Vector Spaces

A *vector space* over some field K is an algebraic structure consisting of a set V on which are defined two algebraic operations: a binary operation referred to as *addition*, and an operation of *multiplication by scalars* in which elements of the vector space are multiplied by elements of the given field K. These two operations are required to satisfy certain axioms that correspond to many of the properties of basic arithmetic. Vector spaces over the field of real numbers are usually referred to as *real vector spaces*. (A real vector space is thus characterized by two operations: an operation in which two elements of the vector space are added together, and an operation in which elements of the vector space are multiplied by real numbers.) Similarly vector spaces over the field of complex numbers are referred to as *complex vector spaces*.

Vector spaces arise in many contexts. A basic example is the vector space consisting of all vectors in 3-dimensional Euclidean space. *Linear algebra*, the algebra of vector spaces, plays a fundamental role in many branches of pure mathematics. The foundations of quantum mechanics are often presented in terms of linear operators acting on an infinite-dimensional complex vector space: this approach was developed by P. A. M. Dirac. The equations of general relativity are usually expressed in the language of *tensors*, operators between certain vector spaces that are derived from the tangent spaces at points of a curved 4-dimensional space-time. In mathematical economics real vector spaces occur naturally when modelling the manufacture of commodities and the exchange of commodities in markets. The study of vector spaces over certain finite fields plays an important role in the design of *error-correcting codes*, which may be used when sending messages along a telephone line, or when one wishes to assign telephone numbers (or analogous identification numbers) in a way that allows for the automatic correction of simple errors occurring if, say, a single digit is incorrectly typed, or if two adjacent digits are transposed.

9.1 The Definition of a Vector Space

Definition. Let K be a field. A *vector space* over the field K consists of a set V on which is defined an operation of *addition* (usually denoted by $+$), associating to elements $\mathbf{u}$ and $\mathbf{v}$ of V an element $\mathbf{u}+\mathbf{v}$ of V, and an operation of *multiplication by scalars*, associating to each element c of K and to each element $\mathbf{v}$ of V an element $c\mathbf{v}$ of V, where the following axioms are satisfied:

- $\mathbf{u}+\mathbf{v}=\mathbf{v}+\mathbf{u}$ for all elements $\mathbf{u}$ and $\mathbf{v}$ of V (i.e., vector addition is *commutative*);
- $(\mathbf{u}+\mathbf{v})+\mathbf{w}=\mathbf{u}+(\mathbf{v}+\mathbf{w})$ for all elements $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ of V (i.e., vector addition is *associative*);
- there exists an an element $\mathbf{0}$ of V (known as the *zero element*) with the property that $\mathbf{v}+\mathbf{0}=\mathbf{v}$ for all elements $\mathbf{v}$ of V;

- given any element $\mathbf{v}$ of V, there exists an element $-\mathbf{v}$ of V with the property that $\mathbf{v}+(-\mathbf{v}) = \mathbf{0}$;
- $(c+d)\mathbf{v} = c\mathbf{v} + d\mathbf{v}$ for all elements c and d of K and elements $\mathbf{v}$ of V;
- $c(\mathbf{v}+\mathbf{w}) = c\mathbf{v} + c\mathbf{w}$ for all elements c of K and elements $\mathbf{v}$ and $\mathbf{w}$ of V;
- $c(d\mathbf{v}) = (cd)\mathbf{v}$ for all elements c and d of K and elements $\mathbf{v}$ of V;
- $1\mathbf{v} = \mathbf{v}$ for all elements $\mathbf{v}$ of V, where 1 is the multiplicative identity element of the field K.

The first four of these axioms (the axioms that involve only the operation of addition) can be summarized in the statement that a vector space is an Abelian group (i.e., a commutative group) with respect to the operation of addition.

Given a vector space V over a field K, we shall refer to the elements of the field K as *scalars*. The scalars of a real vector space are real numbers, and the scalars of a complex vector space are complex numbers. Given an element $\mathbf{v}$ of the vector space V, we shall refer to elements of V that are of the form $c\mathbf{v}$ for some scalar c as *scalar multiples* of $\mathbf{v}$.

A vector space V over a field K is said to be *trivial* if it consists of a single element (which must then be the zero element of V). A vector space with more than one element is said to be *non-trivial*.

9.2 Examples of Vector Spaces

Example. The set of all vectors in 3-dimensional Euclidean space is a real vector space: the vector space axioms in this case are familiar properties of vector algebra.

Example. Given any positive integer n, the set $\mathbb{R}^n$ of all ordered n-tuples $(x_1, x_2, \ldots, x_n)$ of real numbers is a real vector space. Operations of addition of such n-tuples and multiplication of n-tuples by real numbers are defined in the obvious fashion:

$$(x_1, x_2, \ldots, x_n) + (y_1, y_2, \ldots, y_n) = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n),$$

$$c(x_1, x_2, \ldots, x_n) = (cx_1, cx_2, \ldots, cx_n)$$

for all n-tuples $(x_1, x_2, \ldots, x_n)$ and $(y_1, y_2, \ldots, y_n)$ and for all real numbers c. The space $\mathbb{R}^n$ is the natural n-dimensional generalization of the space $\mathbb{R}^3$ of all 3-dimensional vectors, where such vectors are represented with respect to Cartesian coordinates as ordered triples (u, v, w) of real numbers.

Example. Given any positive integer n, the set $\mathbb{C}^n$ of all ordered n-tuples $(z_1, z_2, \ldots, z_n)$ of complex numbers is a complex vector space. The algebraic operations of addition of complex n-tuples and multiplication of complex n-tuples by complex numbers are defined in the obvious fashion, generalizing the corresponding operations on the real vector space $\mathbb{R}^n$.

Example. Let K be any field. Then the set K^n of n-tuples of elements of K is a field over K, where

$$(x_1, x_2, \ldots, x_n) + (y_1, y_2, \ldots, y_n) = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n),$$

$$c(x_1, x_2, \ldots, x_n) = (cx_1, cx_2, \ldots, cx_n)$$

for all elements $(x_1, x_2, \ldots, x_n)$ and $(y_1, y_2, \ldots, y_n)$ of K^n and for all elements c of K. This example is a generalization of the previous two examples.

Example. The set of all polynomials with real coefficients is a real vector space, with the usual operations of addition of polynomials and multiplication of polynomials by scalars (in which all coefficients of the polynomial are multiplied by the same real number). It is easy to verify that the vector space axioms are all satisfied.

Example. The field $\mathbb{Q}(\sqrt{2})$ consisting of all real numbers of the form $p + q\sqrt{2}$, where p and q are required to be rational numbers, is a vector space over the field $\mathbb{Q}$ of rational numbers. The sum of any two numbers in $\mathbb{Q}(\sqrt{2})$ itself belongs to $\mathbb{Q}(\sqrt{2})$, as does the product of a rational number and an number in $\mathbb{Q}(\sqrt{2})$.

Example. The set $C(D, \mathbb{R})$ of all continuous real-valued functions defined over a given subset D of the real numbers is a real vector space: if $x \mapsto f(x)$ and $x \mapsto g(x)$ are continuous functions on D then so are $x \mapsto f(x) + g(x)$ and $x \mapsto cf(x)$ for all real numbers c; moreover these operations of addition of functions and of multiplication of functions by real numbers satisfy the vector space axioms.

Example. The field $\mathbb{C}$ of complex numbers can be viewed as a real vector space: the vector space axioms are satisfied when two complex numbers are added together in the normal fashion, and when complex numbers are multiplied by real numbers.

9.3 Basic Consequences of the Vector Space Axioms

Let V be a vector space over some field K. Then the operation $+$ of addition of elements of V is required to satisfy the following axioms:

- $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$ for all elements $\mathbf{u}$ and $\mathbf{v}$ of V (i.e., vector addition is *commutative*);
- $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$ for all elements $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ of V (i.e., vector addition is *associative*);
- there exists an an element $\mathbf{0}$ of V (known as the *zero element*) with the property that $\mathbf{v} + \mathbf{0} = \mathbf{v}$ for all elements $\mathbf{v}$ of V;
- given any element $\mathbf{v}$ of V, there exists an element $-\mathbf{v}$ of V with the property that $\mathbf{v} + (-\mathbf{v}) = \mathbf{0}$;

These are the axioms that characterize Abelian groups.

We now consider some of the consequences of these axioms. (Corresponding results hold in any Abelian group.)

Lemma 9.1. *Let* $\mathbf{u}$ *and* $\mathbf{v}$ *be elements of a vector space* V*. Then there exists a unique element* $\mathbf{x}$ *of* V *satisfying* $\mathbf{x} + \mathbf{v} = \mathbf{u}$*.*

Proof. The vector space axioms ensure the existence of an element $-\mathbf{v}$ of V with the property that $\mathbf{v} + (-\mathbf{v}) = \mathbf{0}$, where $\mathbf{0}$ is the zero element of V. The identity $\mathbf{x} + \mathbf{v} = \mathbf{u}$ is satisfied when $\mathbf{x} = \mathbf{u} + (-\mathbf{v})$, since

$$(\mathbf{u} + (-\mathbf{v})) + \mathbf{v} = \mathbf{u} + ((-\mathbf{v}) + \mathbf{v}) = \mathbf{u} + (\mathbf{v} + (-\mathbf{v})) = \mathbf{u} + \mathbf{0} = \mathbf{u}.$$

(Here we have used the fact that vector addition is required to be both commutative and associative.) If now $\mathbf{x}$ is any element of V satisfying $\mathbf{x} + \mathbf{v} = \mathbf{u}$ then

$$\mathbf{x} = \mathbf{x} + \mathbf{0} = \mathbf{x} + (\mathbf{v} + (-\mathbf{v})) = (\mathbf{x} + \mathbf{v}) + (-\mathbf{v}) = \mathbf{u} + (-\mathbf{v}).$$

This proves that there is exactly one element $\mathbf{x}$ of V satisfying $\mathbf{x} + \mathbf{v} = \mathbf{u}$, and it is given by the formula $\mathbf{x} = \mathbf{u} + (-\mathbf{v})$. ∎

Let $\mathbf{u}$ and $\mathbf{v}$ be elements of a vector space V. We denote by $\mathbf{u}-\mathbf{v}$ the unique element $\mathbf{x}$ of V with the property satisfying $\mathbf{x}+\mathbf{v}=\mathbf{u}$. Note that $\mathbf{u}-\mathbf{v}=\mathbf{u}+(-\mathbf{v})$ for all elements $\mathbf{u}$ and $\mathbf{v}$ of V. This defines the operation of *subtraction* on any vector space.

If $\mathbf{x}$ is an element of a vector space V and if there exists at least one element $\mathbf{v}$ for which $\mathbf{v}+\mathbf{x}=\mathbf{v}$ then Lemma 9.1 ensures that $\mathbf{x}=\mathbf{0}$. It follows immediately from this that the zero element of a vector space is uniquely determined.

Lemma 9.1 also ensures that, given any element $\mathbf{v}$ of a vector space V there exists exactly one element $-\mathbf{v}$ of V with the property that $\mathbf{v}+(-\mathbf{v})=\mathbf{0}$.

In addition to the axioms for addition listed above, a vector space is required to satisfy axioms that involve the operation of multiplication by scalars. These axioms are as follows:

- $(c+d)\mathbf{v}=c\mathbf{v}+d\mathbf{v}$ for all elements c and d of K and elements $\mathbf{v}$ of V;
- $c(\mathbf{v}+\mathbf{w})=c\mathbf{v}+c\mathbf{w}$ for all elements c of K and elements $\mathbf{v}$ and $\mathbf{w}$ of V;
- $c(d\mathbf{v})=(cd)\mathbf{v}$ for all elements c and d of K and elements $\mathbf{v}$ of V;
- $1\mathbf{v}=\mathbf{v}$ for all elements $\mathbf{v}$ of V, where 1 is the multiplicative identity element of the field K.

We now discuss some elementary consequences of these axioms.

Lemma 9.2. *Let V be a vector space over a field K. Then $c\mathbf{0}=\mathbf{0}$ and $0\mathbf{v}=\mathbf{0}$ for all elements c of K and elements $\mathbf{v}$ of V.*

Proof. The zero element $\mathbf{0}$ of V satisfies $\mathbf{0}+\mathbf{0}=\mathbf{0}$. Therefore

$$c\mathbf{0}+c\mathbf{0}=c(\mathbf{0}+\mathbf{0})=c\mathbf{0}$$

for any element c of K. The elements $c\mathbf{0}$ and $\mathbf{0}$ of V must therefore be equal to one another, since both are equal to the unique element $\mathbf{x}$ of V that satisfies $\mathbf{x}+c\mathbf{0}=c\mathbf{0}$.

The zero element 0 of the field K satisfies $0+0=0$. Therefore

$$0\mathbf{v}+0\mathbf{v}=(0+0)\mathbf{v}=0\mathbf{v}$$

for any element $\mathbf{v}$ of V. The elements $0\mathbf{v}$ and $\mathbf{0}$ of V must therefore be equal to one another, since both are equal to the unique element $\mathbf{y}$ of V that satisfies $\mathbf{y}+0\mathbf{v}=0\mathbf{v}$. ∎

Lemma 9.3. *Let V be a vector space over a field K. Then $(-c)\mathbf{v}=-(c\mathbf{v})$ and $c(-\mathbf{v})=-(c\mathbf{v})$ for all elements c of K and elements $\mathbf{v}$ of V.*

Proof. $(-c)\mathbf{v}=-(c\mathbf{v})$, since

$$c\mathbf{v}+(-c)\mathbf{v}=(c+(-c))\mathbf{v}=0\mathbf{v}=\mathbf{0}.$$

Also $c(-\mathbf{v})=-(c\mathbf{v})$, since

$$c\mathbf{v}+c(-\mathbf{v})=c(\mathbf{v}+(-\mathbf{v}))=c\mathbf{0}=\mathbf{0}.$$ ∎

Lemma 9.4. *Let V be a vector space over a field K. Then $\mathbf{u}-\mathbf{v}=\mathbf{u}+(-1)\mathbf{v}$ for all elements $\mathbf{u}$ and $\mathbf{v}$ of V.*

Proof. The vector space axioms require that $\mathbf{v}=1\mathbf{v}$. It follows from Lemma 9.3 that $-\mathbf{v}=(-1)\mathbf{v}$. Therefore $\mathbf{u}-\mathbf{v}=\mathbf{u}+(-\mathbf{v})=\mathbf{u}+(-1)\mathbf{v}$, as required. ∎

Lemma 9.5. *Let V be a vector space over a field K, let c be an element of K and let $\mathbf{v}$ be an element of V. Suppose that $c\mathbf{v} = \mathbf{0}$. Then either $c = 0$ or $\mathbf{v} = \mathbf{0}$.*

Proof. Suppose that $c\mathbf{v} = \mathbf{0}$ and $c \neq 0$. We must show that $\mathbf{v} = 0$. Now there exists an element c^{-1} of K satisfying $c^{-1}c = 1$, since any non-zero element of a field has a multiplicative inverse. It then follows from the vector space axioms and Lemma 9.2 that

$$\mathbf{v} = 1\mathbf{v} = (c^{-1}c)\mathbf{v} = c^{-1}(c\mathbf{v}) = c^{-1}\mathbf{0} = \mathbf{0},$$

as required. ∎

9.4 Subspaces

Definition. Let V be a vector space over a field K. A non-empty subset U of V is said to be a *subspace* of V if $\mathbf{u} + \mathbf{v} \in U$ and $c\mathbf{u} \in U$ for all elements $\mathbf{u}$ and $\mathbf{v}$ of U and for all elements c of the field K. (Thus the sum of two elements of a subspace of V is required to be an element of the subspace, as is any scalar multiple of an element of that subspace.)

Example. The set of all vectors that are parallel to a given plane is a subspace of the space of all vectors in 3-dimensional Euclidean space.

Example. For each positive integer n, the set of all polynomials with real coefficients whose degree is less than or equal to n is a subspace of the space of all polynomials with real coefficients. (This follows from the fact that the sum of two polynomials of degree not exceeding n is itself such a polynomial, as is any scalar multiple of a polynomial of degree not exceeding n.)

Example. Let D be a subset of the set $\mathbb{R}$ of real numbers, and let $C(D, \mathbb{R})$ be the real vector space consisting of all continuous real-valued functions on D. Given any subset E of D, the set of all continuous real-valued functions f on D with the property that $f(x) = 0$ for all $x \in E$ is a subspace of the vector space $C(D, \mathbb{R})$: the sum of two functions that take the value zero on E is itself a function taking the value zero on E, as is any constant multiple of such a function.

Example. The set of all differentiable real-valued functions on a given interval is a subspace of the real vector space consisting of all continuous real-valued functions on that interval (with the usual operations of addition of functions and of multiplication of functions by real numbers). Indeed the sum of two differentiable functions is itself a differentiable function, as is any constant multiple of a differentiable function.

Lemma 9.6. *Let V be a vector space over a field K. Then any subspace of V is itself a vector space over K.*

Proof. Let U be a subspace of V. If $\mathbf{v}$ is an element of U then so is $(-1)\mathbf{v}$. Now $\mathbf{u} - \mathbf{v} = \mathbf{u} + (-1)\mathbf{v}$ for all elements $\mathbf{u}$ and $\mathbf{v}$ of U (Lemma 9.4), and the sum of two elements of a subspace is itself an element of that subspace. We conclude that if $\mathbf{u}$ and $\mathbf{v}$ are elements of U then so is $\mathbf{u} - \mathbf{v}$.

We must verify that the vector space axioms are satisfied when elements of U are added together or are multiplied by scalars. The operation of addition on U is commutative and associative. The zero element $\mathbf{0}$ of V must belong to U, since subspaces of V are required to be non-empty and $\mathbf{0} = \mathbf{v} - \mathbf{v}$ for any element $\mathbf{v}$ of U. If $\mathbf{v}$ is an element of U then so is $-\mathbf{v}$, since $-\mathbf{v} = \mathbf{0} - \mathbf{v}$. We can therefore conclude that the subspace U is an Abelian group with respect to the operation of addition. The algebraic operations on U of addition and of multiplication by scalars must clearly satisfy the identities listed in the vector space axioms, since these identities are satisfied by the algebraic operations on the vector space V. We conclude therefore that any subspace of V is itself a vector space over the given field. ∎

9.5 Linear Dependence, Spanning Sets and Bases

Let $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ be elements of some vector space over a given field K. An element specified by an expression of the form $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k$ where $c_1, c_2, \ldots, c_k$ are scalars (i.e., elements of the field K) is said to be a *linear combination* of the elements $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$.

Definition. Let V be a vector space over a field K. Elements $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ of V are said to be *linearly dependent* if there exist scalars $c_1, c_2, \ldots, c_k$, not all zero, such that

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k = \mathbf{0}.$$

Elements $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ are said to be *linearly independent* if they are not linearly dependent. (Thus elements $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ of a vector space V are linearly independent if and only if the only solution of the equation

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k = \mathbf{0}$$

is the trivial solution in which the scalars $c_1, c_2, \ldots, c_k$ are all zero.)

Example. The vectors $(2, 2, 5)$, $(3, 3, 12)$ and $(5, 5, -1)$ are linearly dependent elements of the real vector space $\mathbb{R}^3$, since

$$7(2, 2, 5) - 3(3, 3, 12) - (5, 5, -1) = (0, 0, 0).$$

(Thus if $\mathbf{v}_1 = (2, 2, 5)$, $\mathbf{v}_2 = (3, 3, 12)$ and $\mathbf{v}_3 = (5, 5, -1)$, then the equation $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 = \mathbf{0}$ is satisfied with $c_1 = 7$, $c_2 = -3$ and $c_3 = -1$.)

Example. Let $p_j(x) = x^j$ for $j = 0, 1, 2, \ldots, n$, where n is some positive integer. Then the polynomials $p_0(x), p_1(x), p_2(x), \ldots, p_n(x)$ are linearly independent elements of the vector space consisting of all polynomials with real coefficients. Indeed if $c_0, c_1, \ldots, c_n$ are real numbers and if

$$c_0p_0(x) + c_1p_1(x) + c_2p_2(x) + \cdots + c_np_n(x) = 0$$

then $c_0 + c_1x + c_2x^2 + \cdots + c_nx^n$ is the zero polynomial, and therefore $c_j = 0$ for $j = 0, 1, 2, \ldots, n$.

Definition. Elements $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ of a vector space V are said to *span* V if, given any element $\mathbf{u}$ of V, there exist scalars $c_1, c_2, \ldots, c_n$ such that

$$\mathbf{u} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n.$$

Definition. Elements $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ of a vector space V are said to constitute a *basis* of V if these elements are linearly independent and span V.

Example. The vectors $(1, 0, 0)$, $(0, 1, 0)$ and $(0, 0, 1)$ constitute a basis of the vector space $\mathbb{R}^3$ of all ordered triples of real numbers.

Example. Let $\mathbb{R}^n$ be the real vector space consisting of all ordered n-tuples $(x_1, x_2, \ldots, x_n)$ of real numbers, and, for each integer j between 1 and n, let $\mathbf{e}_j$ be the ordered n-tuple whose jth component is equal to one and whose other components are zero. Then

$$(x_1, x_2, \ldots, x_n) = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_n\mathbf{e}_n$$

for any ordered n-tuple $(x_1, x_2, \ldots, x_n)$ of real numbers. It follows directly from this that the elements $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ are linearly independent and span the vector space $\mathbb{R}^n$. Thus $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ is a basis of $\mathbb{R}^n$.

Example. Let $p_j(x) = x^j$ for $j = 0, 1, 2, \ldots, n$, where n is some positive integer. Then the linearly independent polynomials $p_0(x), p_1(x), p_2(x), \ldots, p_n(x)$ span the vector space consisting of all polynomials with real coefficients whose degree does not exceed n, since

$$c_0 + c_1x + c_2x^2 + \cdots + c_nx^n = c_0p_0(x) + c_1p_1(x) + c_2p_2(x) + \cdots + c_np_n(x)$$

for all polynomials $c_0 + c_1x + c_2x^2 + \cdots + c_nx^n$ with real coefficients. We conclude that $1, x, x^2, \ldots, x^n$ is a basis of the vector space consisting of all polynomials with real coefficients whose degree does not exceed n.

Theorem 9.7. *Elements $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ of a vector space V constitute a basis of that vector space if and only if, given any element $\mathbf{u}$ of V, there exist uniquely determined scalars $c_1, c_2, \ldots, c_n$ such that*

$$\mathbf{u} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n.$$

Proof. First suppose that $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ is a list of elements of V with the property that, given any element $\mathbf{u}$ of V, there exist uniquely determined scalars $c_1, c_2, \ldots, c_n$ such that

$$\mathbf{u} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n.$$

Then the elements $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ span V. Also the uniqueness of the scalars ensures that the zero element $\mathbf{0}$ of V cannot be expressed as a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ unless the scalars involved are all zero. Therefore these elements are linearly independent and thus constitute a basis of the vector space V.

Conversely suppose that $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ is a basis of V. Then any element of V can be expressed as a linear combination of the basis vectors. We must prove that the scalars involved are uniquely determined. Let $c_1, c_2, \ldots, c_n$ and $d_1, d_2, \ldots, d_n$ be scalars satisfying

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n = d_1\mathbf{v}_1 + d_2\mathbf{v}_2 + \cdots + d_n\mathbf{v}_n.$$

Then

$$(c_1 - d_1)\mathbf{v}_1 + (c_2 - d_2)\mathbf{v}_2 + \cdots + (c_n - d_n)\mathbf{v}_n = \mathbf{0}.$$

But then $c_j - d_j = 0$ and thus $c_j = d_j$ for $j = 1, 2, \ldots, n$, since the elements of any basis are required to be linearly independent. This proves that any element of V can be represented in a unique fashion as a linear combination of the elements of a basis of V, as required. ∎

9.6 Finite-Dimensional Vector Spaces

Definition. A vector space V is said to be *finite-dimensional* if there exists a finite subset of V whose elements span V.

A vector space is said to be trivial if it consists of a single element (the zero element). We shall show that every non-trivial finite-dimensional vector space has a basis (Corollary 9.10). Moreover any two bases of a finite-dimensional vector space have the same number of elements (Corollary 9.13). This enables us to define the *dimension* of a non-trivial finite-dimensional vector space to be the number of elements in any basis of that vector space. The dimension of a trivial vector space is defined to be zero. Any subspace of a finite-dimensional vector space V is itself a finite-dimensional vector space whose dimension does not exceed that of V (Proposition 9.14).

Proposition 9.8. *Let V be a non-trivial vector space, and let S be a finite subset of V whose elements span V. Let n be the smallest positive integer for which there exists a set of n elements of S that span V. Then any n vectors of S that span V are linearly independent, and thus constitute a basis of V.*

Proof. Let $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ be n elements of S which span V. We show that these elements are linearly independent.

Suppose that $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ were linearly dependent. Then $n > 1$, and there would exist scalars $a_1, a_2, \ldots, a_n$, not all zero, such that

$$a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_n\mathbf{v}_n = \mathbf{0}.$$

We may suppose, without loss of generality, that $a_n \neq 0$. Then

$$\mathbf{v}_n = b_1\mathbf{v}_1 + b_2\mathbf{v}_2 + \cdots + b_{n-1}\mathbf{v}_{n-1},$$

where $b_i = -a_i a_n^{-1}$ for $i = 1, 2, \ldots, n-1$. But then $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{n-1}$ would span V, since any linear combination of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ could be expressed as a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{n-1}$. (Indeed

$$\sum_{i=1}^{n} c_i\mathbf{v}_i = \sum_{i=1}^{n-1}(c_i + c_n b_i)\mathbf{v}_i$$

for all scalars $c_1, c_2, \ldots, c_n$.) But the definition of n ensures that no set of $n-1$ elements of S can span V. We conclude that $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ must be linearly independent, and thus must constitute a basis of V, as required. ∎

Corollary 9.9. *Let V be a non-trivial vector space, and let S be a finite subset of V whose elements span V. Then there exists a basis of V whose elements belong to S.*

Corollary 9.10. *Every non-trivial finite-dimensional vector space has a basis.*

Proposition 9.11. *Let V be a non-trivial vector space, let S be a finite subset of V whose elements span V, and let $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m$ be linearly independent elements of S. Let n be the smallest positive integer for which there exists a set of n elements of S that span V. Then $m \leq n$, and the elements $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m$ can be included in a basis of V that consists of n elements of S.*

Proof. We claim that if k is a non-negative integer less than m and n, and if the elements $\mathbf{x}_i$ for $i \leq k$ can be included in a basis of V consisting of n elements of S, then so can the elements $\mathbf{x}_i$ for $i \leq k+1$. Suppose therefore that $\mathbf{v}_1, \ldots, \mathbf{v}_{n-k}$ are elements of S and that the elements $\mathbf{x}_i$ for $1 \leq i \leq k$ together with the elements $\mathbf{v}_i$ for $1 \leq i \leq n-k$ constitute a basis of V. Then the elements $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{k+1}, \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{n-k}$ are linearly dependent, since $\mathbf{x}_{k+1}$ can be expressed as a linear combination of the other elements in this list. Thus there exist scalars $a_1, a_2, \ldots, a_{k+1}$ and $b_1, b_2, \ldots, b_{n-k}$, not all zero, such that

$$\sum_{i=1}^{k+1} a_i\mathbf{x}_i + \sum_{i=1}^{n-k} b_i\mathbf{v}_i = \mathbf{0}.$$

The linear independence of $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{k+1}$ ensures that the scalars $b_1, b_2, \ldots, b_{n-k}$ are not all zero. Without loss of generality we may suppose that $b_{n-k} \neq 0$. Then

$$\mathbf{v}_{n-k} = -\sum_{i=1}^{k+1} a_i b_{n-k}^{-1}\mathbf{x}_i - \sum_{i=1}^{n-k-1} b_i b_{n-k}^{-1}\mathbf{v}_i.$$

It follows from this that V is spanned by the n elements $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{k+1}$ and $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{n-k-1}$. But n is the smallest positive integer for which there exist n elements of S that span V. It follows from Proposition 9.8 that any n vectors of S that span V constitute a basis of V. Therefore the elements $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{k+1}$ and $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{n-k-1}$ constitute a basis of V.

We have shown that if k is a non-negative integer less than m and n, and if the elements $\mathbf{x}_i$ with $i \leq k$ can be included in a basis of V consisting of n elements of S, then so can the elements $\mathbf{x}_i$ with $i \leq k+1$. Given any positive integer j that does not exceed m or n we can apply this result with $k = 0, 1, \ldots, j-1$. We deduce that the elements $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_j$ can be included in a basis of V consisting of n elements of S, provided that j does not exceed m or n.

Suppose it were the case that $m > n$. Then it would follow (on taking $j = n$) that the elements $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ would constitute a basis of V, and therefore each of the vectors $\mathbf{x}_{n+1}, \ldots, \mathbf{x}_m$ would be expressed as a linear combination of $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$. But this would contradict the linear independence of $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m$. Therefore $m \leq n$. It then follows (on taking $j = m$) that $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m$ can be included in a basis of V consisting of n elements of S, as required. ∎

Corollary 9.12. *Let V be a non-trivial vector space, and let X and Y be finite subsets of V. Suppose that the elements of X are linearly independent and the elements of Y span V. Then the number of elements of X does not exceed the number of elements of Y.*

Proof. Let r and s denote the number of elements in X and Y respectively, and let $S = X \cup Y$. Then the elements of S span V. Let n be the smallest positive integer for which there exists a set of n elements of S that span V. It follows from Proposition 9.11 that $r \leq n$. But $n \leq s$, since the elements of Y span V. Therefore $r \leq s$, as required. ∎

Corollary 9.13. *Any two bases of a finite-dimensional vector space contain the same number of elements.*

Proof. This result follows immediately from Corollary 9.13, since the elements of any basis of a finite-dimensional vector space are linearly independent and span the vector space. ∎

Definition. The *dimension* of a finite-dimensional vector space is defined to be number of elements in any basis of that vector space. The dimension is defined to be zero in the case where the vector space consists of just the zero element.

Example. The vector space $\mathbb{R}^n$ consisting of all n-tuples of real numbers is an n-dimensional real vector space.

Example. The field $\mathbb{C}$ of complex numbers is a 2-dimensional real vector space: the numbers 1 and $\sqrt{-1}$ constitute a basis of $\mathbb{C}$ as a real vector space since any complex number can be expressed uniquely in the form $x + y\sqrt{-1}$, where x and y are required to be real numbers.

Example. Let n be a positive integer. The vector space consisting of all polynomials with real coefficients whose degree does not exceed n is an $(n+1)$-dimensional real vector space: the polynomials $1, x, x^2, \ldots, x^n$ constitute a basis for this vector space.

Proposition 9.14. *A subspace U of a finite-dimensional vector space V is itself a finite-dimensional vector space whose dimension cannot exceed that of V. Moreover if m and n are the dimensions of U and V then there exists a basis $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ of V with the property that the first m elements of this basis constitute a basis of the subspace U.*

Proof. Let U be a subspace of a finite-dimensional vector space V. The result is trivial when $U = \{\mathbf{0}\}$. Suppose then that U is non-trivial. Now Corollary 9.12 ensures that the number of linearly independent elements in any subset of a finite-dimensional vector space V cannot exceed the dimension of V. Let m be the largest number of linearly independent elements in any subset of U, and let $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ be linearly independent elements of U. We claim that these elements span U and therefore constitute a basis for U.

Let $\mathbf{u}$ be an element of U. Then the elements $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m, \mathbf{u}$ must be linearly dependent, since this list contains $m+1$ members. It follows that there exist scalars $c_1, c_2, \ldots, c_m$ and d, not all zero, such that

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_m\mathbf{v}_m + d\mathbf{u} = \mathbf{0}.$$

The linear independence of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ then ensures that $d \neq 0$. Then

$$\mathbf{u} = -c_1d^{-1}\mathbf{v}_1 - c_2d^{-1}\mathbf{v}_2 - \cdots - c_md^{-1}\mathbf{v}_m.$$

We conclude that the linearly independent elements $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ span U and thus constitute a basis of U. Moreover the dimension m of U does not exceed the dimension n of V.

Finally let $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m\} \cup S_0$, where S_0 is a finite subset of V whose elements span V. It follows from Proposition 9.11 that there exists a basis of V consisting of elements of S which includes $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$. Therefore there exist elements $\mathbf{v}_i$ of S_0 for $m < i \leq n$ such that $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ is a basis of V. The first m elements of this basis constitute a basis of U, as required. ∎

Example. The space consisting of all polynomials with real coefficients is an infinite-dimensional real vector space. For if this space were a finite-dimensional vector space whose dimension is N then no linearly independent subset of the vector space could contain more than N elements. But the polynomials $1, x, x^2, \ldots, x^n$ are linearly independent for each positive integer n. Therefore the space of all polynomials with real coefficients cannot be finite-dimensional.

Example. The space consisting of all continuous real-valued functions defined on the set $\mathbb{R}$ of real numbers is an infinite-dimensional real vector space since the polynomial functions constitute an infinite-dimensional subspace, and a finite-dimensional vector space cannot contain any infinite-dimensional subspace (Proposition 9.14). ∎

9.7 Linear Transformations

Definition. Let V and W be vector spaces over some field K. A function $T: V \to W$ is said to be a *linear transformation* if $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ and $T(c\mathbf{v}) = cT(\mathbf{v})$ for all elements $\mathbf{u}$ and $\mathbf{v}$ of V and for all elements c of K.

Let V and W be vector spaces over a field K, and let $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ be elements of a vector space V, and let $T: V \to W$ be a linear transformation from V to W. Let $\mathbf{w}_j = T(\mathbf{v}_j)$ for $j = 1, 2, \ldots, n$. Then

$$\begin{aligned} T(c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n) &= T(c_1\mathbf{v}_1) + T(c_2\mathbf{v}_2) + \cdots + T(c_n\mathbf{v}_n) \\ &= c_1\mathbf{w}_1 + c_2\mathbf{w}_2 + \cdots + c_n\mathbf{w}_n \end{aligned}$$

for all scalars $c_1, c_2, \ldots, c_n$.

In particular, $\mathbb{R}^n$ be the space of all ordered n-tuples $(x_1, x_2, \ldots, x_n)$ of real numbers, and, for each integer j between 1 and n, let $\mathbf{e}_j$ be the ordered n-tuple whose jth component is equal to one and whose other components are zero. Now

$$(x_1, x_2, \ldots, x_n) = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_n\mathbf{e}_n$$

for any ordered n-tuple $(x_1, x_2, \ldots, x_n)$. Thus if $T: \mathbb{R}^n \to W$ is a linear transformation from $\mathbb{R}^n$ to some real vector space W then

$$T(x_1, x_2, \ldots, x_n) = x_1\mathbf{w}_1 + x_2\mathbf{w}_2 + \cdots + x_n\mathbf{w}_n,$$

where $\mathbf{w}_j = T(\mathbf{e}_j)$ for $j = 1, 2, \ldots, n$.

Lemma 9.15. *Let V and W be vector spaces over a given field, and let $T: V \to W$ be a linear transformation from V to W. Then $T(\mathbf{u} - \mathbf{v}) = T(\mathbf{u}) - T(\mathbf{v})$ for all elements $\mathbf{u}$ and $\mathbf{v}$ of V, and $T(\mathbf{0}) = \mathbf{0}$.*

Proof. Let $\mathbf{u}$ and $\mathbf{v}$ be elements of V. Then $\mathbf{u} - \mathbf{v} = \mathbf{u} + (-1)\mathbf{v}$ (Lemma 9.4), and hence

$$T(\mathbf{u} - \mathbf{v}) = T(\mathbf{u} + (-1)\mathbf{v}) = T(\mathbf{u}) + (-1)T(\mathbf{v}) = T(\mathbf{u}) - T(\mathbf{v}).$$

In particular $T(\mathbf{0}) = T(\mathbf{v} - \mathbf{v}) = T(\mathbf{v}) - T(\mathbf{v}) = \mathbf{0}$, as required. ∎

Definition. The *kernel* $\ker T$ of a linear transformation $T: V \to W$ is defined by

$$\ker T = \{\mathbf{v} \in V : T(\mathbf{v}) = \mathbf{0}\}.$$

Lemma 9.16. *The kernel $\ker T$ of a linear transformation $T: V \to W$ is a subspace of V.*

Proof. The kernel $\ker T$ is non-empty, since $\mathbf{0} \in \ker T$. Let $\mathbf{u}$ and $\mathbf{v}$ be elements of $\ker T$. Then

$$T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v}) = \mathbf{0} + \mathbf{0} = \mathbf{0},$$

and therefore $\mathbf{u} + \mathbf{v}$ is also an element of $\ker T$. Moreover if $\mathbf{v}$ is an element of $\ker T$ then so is $c\mathbf{v}$, since $T(c\mathbf{v}) = cT(\mathbf{v}) = c\mathbf{0} = \mathbf{0}$. Thus $\ker T$ is a subspace of V. ∎

Definition. The *range* (or *image*) $T(V)$ of a linear transformation $T: V \to W$ is defined by

$$T(V) = \{T(\mathbf{v}) : \mathbf{v} \in V\}.$$

Lemma 9.17. *The range $T(V)$ of a linear transformation $T: V \to W$ is a subspace of W.*

Proof. The range $T(V)$ is clearly non-empty. The sum of two elements of $T(V)$ must belong to $T(V)$ since $T(\mathbf{u}) + T(\mathbf{v}) = T(\mathbf{u} + \mathbf{v})$ for all elements $\mathbf{u}$ and $\mathbf{v}$ of V. Also any scalar multiple of an element of $T(V)$ must belong to $T(V)$, since $cT(\mathbf{v}) = T(c\mathbf{v})$ for all scalars c and elements $\mathbf{v}$ of V. ∎

9.8 Representation of Linear Transformations by Matrices

Let V and W be vector spaces of dimensions m and n respectively over some field K, and let $T: V \to W$ be a linear transformation from V to W. Let $\mathbf{v}_1, \mathbf{v}_2, \ldots \mathbf{v}_m$ be a basis for V, and let $\mathbf{w}_1, \mathbf{w}_2, \ldots \mathbf{w}_n$ be a basis for W. An element $\mathbf{u}$ of V can then be expressed as a linear combination of elements of the given basis of V:

$$\mathbf{u} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_m\mathbf{v}_m,$$

where $x_1, x_2, \ldots, x_m$ are scalars (i.e., elements of the field K). Similarly the image $T(\mathbf{u})$ of $\mathbf{u}$ under the linear transformation T can also be expressed as a linear combination of elements of the given basis of W:

$$T(\mathbf{u}) = y_1\mathbf{w}_1 + y_2\mathbf{w}_2 + \cdots + y_n\mathbf{w}_n,$$

where $y_1, y_2, \ldots, y_n$ are scalars. But if $T: V \to W$ is a linear transformation then

$$\begin{aligned} T(\mathbf{u}) &= T\left(x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_m\mathbf{v}_m\right) \\ &= T(x_1\mathbf{v}_1) + T(x_2\mathbf{v}_2) + \cdots + T(x_m\mathbf{v}_m) \\ &= x_1T(\mathbf{v}_1) + x_2T(\mathbf{v}_2) + \cdots + x_mT(\mathbf{v}_m). \end{aligned}$$

Moreover $T(\mathbf{v}_1), T(\mathbf{v}_2), \ldots T(\mathbf{v}_m)$ are elements of the vector space W, and we can therefore write $T(\mathbf{v}_k) = \sum_{j=1}^{n} M_{jk}\mathbf{w}_j$, where the quantities M_{jk} are scalars. It follows that

$$T(\mathbf{u}) = T\left(\sum_{k=1}^{m} x_k\mathbf{v}_k\right) = \sum_{k=1}^{m} x_kT(\mathbf{v}_k) = \sum_{k=1}^{m} x_k\left(\sum_{j=1}^{n} M_{jk}\mathbf{w}_j\right) = \sum_{j=1}^{n}\sum_{k=1}^{m} M_{jk}x_k\mathbf{w}_j.$$

Examination of this formula shows that

$$y_j = \sum_{k=1}^{m} M_{jk}x_k.$$

(Thus $y_1 = M_{11}x_1 + M_{12}x_2 + \cdots + M_{1m}x_m$ etc.) The relation between the coefficients $x_1, x_2, \ldots, x_m$ and $y_1, y_2, \ldots, y_n$ can be expressed more succinctly by the matrix equation $\mathbf{y} = M\mathbf{x}$, where $\mathbf{x}$ and $\mathbf{y}$ are the *column vectors* given by

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \text{ and } \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}.$$

and M is the matrix with the value M_{jk} in the jth row and kth column. For example, suppose that the dimensions of V and W are 3 and 2 respectively. Then

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} M_{11} & M_{12} & M_{13} \\ M_{21} & M_{22} & M_{23} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix},$$

where the coefficients M_{jk} of the 2×3 matrix representing the linear transformation T are determined so that

$$\begin{aligned} T(\mathbf{v}_1) &= M_{11}\mathbf{w}_1 + M_{21}\mathbf{w}_2 \\ T(\mathbf{v}_2) &= M_{12}\mathbf{w}_1 + M_{22}\mathbf{w}_2 \\ T(\mathbf{v}_3) &= M_{13}\mathbf{w}_1 + M_{23}\mathbf{w}_2 \end{aligned}$$

Example. An anticlockwise rotation about the vertical axis though an angle of θ radians sends a vector with Cartesian components (u, v, w) to the vector (u', v', w'), where

$$u' = u\cos\theta - v\sin\theta, \qquad v' = u\sin\theta + v\cos\theta, \qquad w' = w.$$

This rotation is thus a linear transformation on the space $\mathbb{R}^3$ of vectors in 3-dimensional Euclidean space. The three vectors $(1, 0, 0)$, $(0, 1, 0)$ and $(0, 0, 1)$ constitute a basis of $\mathbb{R}^3$ and the rotation is represented with respect to this basis by the 3×3 matrix

$$\begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Example. Let n be a positive integer, let V_n be the real vector space consisting of all polynomials with real coefficients whose degree does not exceed n, and let V_{n-1} be the subspace of V_n consisting of those polynomials whose degree does not exceed $n-1$. (The algebraic operations of addition and of multiplication by real numbers are defined on V_n in the usual fashion.) It is easily seen that the function that sends a polynomial $p(x)$ to its derivative $p'(x)$ is a linear transformation from V_n to V_{n-1}, and therefore is represented by a matrix with respect to chosen bases of V_n and V_{n-1}.

Suppose for example that we take $n = 3$. We can take as our basis for the space V_3 the four polynomials 1, x, x^2 and x^3. The polynomials 1, x and x^2 will then provide a basis for the space V_2. The matrix of the linear transformation sending a polynomial to its derivative is represented with respect to these bases of V_3 and V_2 by the 3×4 matrix

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}$$

since the derivatives of the polynomials 1, x, x^2 and x^3 constituting the chosen basis of V_3 are 0, 1, $2x$ and $3x^2$ respectively.

Let U, V and W be vector spaces over a field K, let $\mathbf{u}_1, \mathbf{u}_2, \ldots \mathbf{u}_l$ be a basis for U, let $\mathbf{v}_1, \mathbf{v}_2, \ldots \mathbf{v}_m$ be a basis for V, and let $\mathbf{w}_1, \mathbf{w}_2, \ldots \mathbf{w}_n$ be a basis for W. Let $S: U \to V$ and $T: V \to W$ be linear transformations, and let L and M be the matrices representing the linear transformations S and T respectively with respect to the chosen bases of U, V and W. Then $S(\mathbf{u}_k) = \sum_{j=1}^{m} L_{jk}\mathbf{v}_j$ and $T(\mathbf{v}_j) = \sum_{i=1}^{n} M_{ij}\mathbf{w}_i$, and hence

$$\begin{aligned} TS(\mathbf{u}_k) &= T\left(\sum_{j=1}^{m} L_{jk}\mathbf{v}_j\right) = \sum_{j=1}^{m} L_{jk}T(\mathbf{v}_j) = \sum_{j=1}^{m} L_{jk}\left(\sum_{i=1}^{n} M_{ij}\mathbf{w}_i\right) \\ &= \sum_{i=1}^{n}\left(\sum_{j=1}^{m} M_{ij}L_{jk}\right)\mathbf{w}_i = \sum_{i=1}^{n}(ML)_{ik}\mathbf{w}_i, \end{aligned}$$

where $(ML)_{ik}$ denotes the element in the ith row and kth column of the product matrix ML. This calculation demonstrates that when linear transformations are represented by matrices with respect to chosen bases of the vector spaces involved, the composition of two linear transformations is represented by the product of the corresponding matrices.

Proposition 9.18. *Let V and W be finite-dimensional vector spaces over some given field, and let $T: V \to W$ be a linear transformation from V to W. Let $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_r$ be a basis of $T(V)$, and let $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ be elements of V, where $n \geq r$. Suppose that $T(\mathbf{v}_j) = \mathbf{w}_j$ for $j = 1, 2, \ldots, r$. If $\ker T = \{\mathbf{0}\}$ suppose that $n = r$. If $\ker T \neq \{\mathbf{0}\}$ suppose that $n > r$ and that the elements $\mathbf{v}_j$ with $j > r$ constitute a basis for the kernel $\ker T$ of T. Then $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ is a basis of V.*

Proof. Let $\mathbf{u}$ be an element of V. Then $T(\mathbf{u})$ belongs to the range $T(V)$ of T, and hence there exist scalars $c_1, c_2, \ldots, c_r$ such that $T(\mathbf{u}) = \sum_{j=1}^{r} c_j\mathbf{w}_j$. Then

$$T\left(\mathbf{u} - \sum_{j=1}^{r} c_j\mathbf{v}_j\right) = T(\mathbf{u}) - \sum_{j=1}^{r} c_j\mathbf{w}_j = \mathbf{0}.$$

and thus $\mathbf{u} - \sum_{j=1}^{r} c_j\mathbf{v}_j$ belongs to the kernel of T. If $n = r$ then the kernel of T consists of just the zero vector, and therefore $\mathbf{u} = \sum_{j=1}^{r} c_j\mathbf{v}_j$. If on the other hand $n > r$ then any element of $\ker T$ can

be expressed as a linear combination of the elements $\mathbf{v}_{r+1}, \ldots, \mathbf{v}_n$, since these elements constitute a basis of $\ker T$, and therefore there exist scalars $c_{r+1}, \ldots, c_n$ such that

$$\mathbf{u} - \sum_{j=1}^{r} c_j \mathbf{v}_j = \sum_{j=r+1}^{n} c_j \mathbf{v}_j.$$

Then $\mathbf{u} = \sum_{j=1}^{n} c_j \mathbf{v}_j$. We conclude that the elements $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ span the vector space V.

We now show that the elements $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ are linearly independent. Let $c_1, c_2, \ldots, c_n$ be scalars satisfying $\sum_{j=1}^{n} c_j \mathbf{v}_j = \mathbf{0}$. Now

$$\sum_{j=1}^{r} c_j \mathbf{w}_j = \sum_{j=1}^{n} c_j T(\mathbf{v}_j) = T\left(\sum_{j=1}^{n} c_j \mathbf{v}_j\right) = \mathbf{0},$$

since $T(\mathbf{v}_j) = \mathbf{w}_j$ if $1 \leq j \leq r$ and $T(\mathbf{v}_j) = \mathbf{0}$ if $r < j \leq n$. But the elements $\mathbf{w}_1, \mathbf{w}_2 \ldots, \mathbf{w}_r$ are linearly independent, since they constitute a basis of $T(V)$. It follows from this that $c_j = 0$ for each j between 1 and r. If $n = r$ then we can conclude immediately that the elements $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ are linearly independent. Suppose that $n > r$. Then $\sum_{j=r+1}^{n} c_j \mathbf{v}_j = \mathbf{0}$, since $\sum_{j=1}^{n} c_j \mathbf{v}_j = \mathbf{0}$ and $c_j = 0$ when $j \leq r$. But the elements $\mathbf{v}_j$ with $r < j \leq n$ are linearly independent, since they constitute a basis for the kernel of T. Thus if $n > r$ then $c_j = 0$ for each j satisfying $r < j \leq n$. We have thus shown that if $\sum_{j=1}^{n} c_j \mathbf{v}_j = \mathbf{0}$ then $c_j = 0$ for each integer j between 1 and n. We conclude that the elements $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ are linearly independent, and thus constitute a basis of V. ∎

Corollary 9.19. *Let V and W be finite-dimensional vector spaces over some given field, and let $T: V \to W$ be a linear transformation from V to W. Let $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_r$ be a basis of $T(V)$. Then there exists a basis $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ of V, where $n \geq r$, such that*

$$T(\mathbf{v}_j) = \begin{cases} \mathbf{w}_j & \text{if } 1 \leq j \leq r; \\ \mathbf{0} & \text{if } r < j \leq n. \end{cases}$$

Moreover if $n > r$ then the elements $\mathbf{v}_j$ with $r < j \leq n$ constitute a basis for the kernel of T.

Proof. The existence of the required basis $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ follows on applying Proposition 9.18. If $\sum_{j=1}^{n} c_j \mathbf{v}_j$ belongs to the kernel of T then $c_j = 0$ for $j = 1, 2, \ldots, r$, since $T(\sum_{j=1}^{n} c_j \mathbf{v}_j) = \sum_{j=1}^{r} c_j \mathbf{w}_j$. It follows that if $n > r$ then the elements $\mathbf{v}_j$ with $r < j \leq n$ span the kernel of T. These elements are also linearly independent. They therefore constitute a basis of the kernel of T, as required. ∎

Let V and W be finite-dimensional vector spaces over some given field, let $T: V \to W$ be a linear transformation from V to W. The *rank* of T is defined to be the dimension of the range $T(V)$ of T, and the *nullity* of T is defined to be the dimension of the kernel of T. The dimension of a vector space is by definition the number of elements in a basis of that vector space. The following result therefore follows immediately from Corollary 9.19.

Corollary 9.20. *Let V and W be finite-dimensional vector spaces over some given field, and let $T: V \to W$ be a linear transformation from V to W. Then the sum of the rank and nullity of T is equal to the dimension of V.*

Example. Let $T: V \to W$ be a linear transformation of rank 2 from a vector space V of dimension 4 to a vector space W of dimension 3. The range $T(V)$ of T is 2-dimensional. We see from Proposition 9.14 that there exists a basis $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3$ of W with the property that the elements $\mathbf{w}_1$ and $\mathbf{w}_2$ belong to

the range $T(V)$ and constitute a basis for $T(V)$. Then Theorem 9.19 shows the existence of a basis $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4$ of V such that $T(\mathbf{v}_1) = \mathbf{w}_1$, $T(\mathbf{v}_2) = \mathbf{w}_2$, $T(\mathbf{v}_3) = \mathbf{0}$ and $T(\mathbf{v}_4) = \mathbf{0}$. The matrix representing the linear transformation $T: V \to W$ with respect to these bases of V and W is then

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

This example can easily be generalized to apply to linear transformations between finite-dimensional vector spaces of arbitrary dimensions.

9.9 Isomorphisms of Vector Spaces

A function $f: X \to Y$ from a set X to a set Y is said to be *injective* if, given any element y of Y, there exists at most one element x of X satisfying $f(x) = y$. A function $f: X \to Y$ is said to be *surjective* if, given any element y of Y, there exists at least one element x of X satisfying $f(x) = y$. A function $f: X \to Y$ is said to be *bijective* if, given any element y of Y, there exists exactly one element x of X satisfying $f(x) = y$. We see from these definitions that a function is bijective if and only if it is both injective and surjective. Injective, surjective and bijective functions are referred to as *injections*, *surjections* and *bijections*.

A function $f^{-1}: Y \to X$ is said to be the *inverse* of a function $f: X \to Y$ if $f^{-1}(f(x)) = x$ for all $x \in X$ and $f(f^{-1}(y)) = y$ for all $y \in Y$. It is a straightforward exercise to show that a function between two sets is bijective if and only if it has a well-defined inverse: the inverse $f^{-1}: Y \to X$ of a bijection $f: X \to Y$ is the function that sends an element y of Y to the unique element x of X satisfying $f(x) = y$. The inverse of a bijection is itself a bijection.

Definition. Let V and W be vector spaces over a given field. An *isomorphism* from V to W is a linear transformation $T: V \to W$ that is also a bijection from V to W. The vector spaces V and W are said to be *isomorphic* if there exists an isomorphism from V to W.

We recall that the kernel of a linear transformation $T: V \to W$ between vector spaces V and W is the subspace of V consisting of all vectors of V that are sent by T to the zero vector of W.

Lemma 9.21. *A linear transformation $T: V \to W$ is injective if and only if $\ker T = \{\mathbf{0}\}$.*

Proof. Suppose that $T: V \to W$ is injective. Then there can be at most one element $\mathbf{v}$ of V satisfying $T(\mathbf{v}) = \mathbf{0}$. But $T(\mathbf{0}) = \mathbf{0}$ (Lemma 9.15). Therefore $\ker T = \{\mathbf{0}\}$.

Conversely suppose that $T: V \to W$ is a linear transformation with the property that $\ker T = \{\mathbf{0}\}$. Let $\mathbf{u}$ and $\mathbf{v}$ be elements of V satisfying $T(\mathbf{u}) = T(\mathbf{v})$. Then $T(\mathbf{u} - \mathbf{v}) = T(\mathbf{u}) - T(\mathbf{v}) = \mathbf{0}$ and hence $\mathbf{u} - \mathbf{v} \in \ker T$. But then $\mathbf{u} - \mathbf{v} = \mathbf{0}$, and hence $\mathbf{u} = \mathbf{v}$. Thus if $\ker T = \{\mathbf{0}\}$ then $T: V \to W$ is injective, as required. ∎

Lemma 9.22. *Let V and W be vector spaces over a given field. A linear transformation $T: V \to W$ from V to W is an isomorphism if and only if $\ker T = \{\mathbf{0}\}$ and $T(V) = W$.*

Proof. The result follows immediately from the fact that the linear transformation $T: V \to W$ is injective if and only if $\ker T = \{\mathbf{0}\}$, and $T: V \to W$ is surjective if and only if $T(V) = W$.

Lemma 9.23. *Let V and W be vector spaces over a given field, and let $T: V \to W$ be an isomorphism from V to W. Then the inverse $T^{-1}: W \to V$ of the function T is well-defined and is an isomorphism from W to V.*

Proof. A function is bijective if and only if it has a well-defined inverse. Therefore, given an isomorphism $T: V \to W$, there exists a well-defined function $T^{-1}: W \to V$ that is an inverse of the function $T: V \to W$. Moreover the function $T^{-1}: W \to V$ is a bijection. It remains to show that the inverse function T^{-1} is a linear transformation.

Let $\mathbf{w}$ and $\mathbf{x}$ be elements of W. Then there exist unique elements $\mathbf{u}$ and $\mathbf{v}$ of V satisfying $T(\mathbf{u}) = \mathbf{w}$ and $T(\mathbf{v}) = \mathbf{x}$. Then $T(\mathbf{u} + \mathbf{v}) = \mathbf{w} + \mathbf{x}$, and thus

$$T^{-1}(\mathbf{w} + \mathbf{x}) = \mathbf{u} + \mathbf{v} = T^{-1}(\mathbf{w}) + T^{-1}(\mathbf{x}).$$

Also $T^{-1}(c\mathbf{w}) = c\mathbf{u} = cT^{-1}(\mathbf{w})$ for all scalars c. Thus the bijection $T^{-1}: W \to V$ is indeed a linear transformation, and is therefore an isomorphism from W to V, as required. ∎

Lemma 9.24. *Let V and W be vector spaces over a given field, and let $T: V \to W$ be an isomorphism from V to W. Let $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ be elements of V.*

(i) *If $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ are linearly independent, then so are $T(\mathbf{v}_1), T(\mathbf{v}_2), \ldots, T(\mathbf{v}_n)$.*

(ii) *If $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ span the vector space V, then $T(\mathbf{v}_1), T(\mathbf{v}_2), \ldots, T(\mathbf{v}_n)$ span the vector space W.*

(iii) *If $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ is a basis of V, then $T(\mathbf{v}_1), T(\mathbf{v}_2), \ldots, T(\mathbf{v}_n)$ is a basis of W.*

Proof. Suppose that $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ are linearly independent elements of V. Let $c_1, c_2, \ldots, c_n$ be scalars satisfying $\sum_{j=1}^{n} c_j T(\mathbf{v}_j) = \mathbf{0}$. We must prove that $c_j = 0$ for all j. Now $T\left(\sum_{j=1}^{n} c_j \mathbf{v}_j\right) = \sum_{j=1}^{n} c_j T(\mathbf{v}_j)$, since $T: V \to W$ is a linear transformation. It follows that $\sum_{j=1}^{n} c_j \mathbf{v}_j$ belongs to the kernel of T. But $\ker T = \{\mathbf{0}\}$, since $T: V \to W$ is an isomorphism (Lemma 9.22). Therefore $\sum_{j=1}^{n} c_j \mathbf{v}_j = \mathbf{0}$. It follows now from the linear independence of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ that $c_j = 0$ for all j. Thus $T(\mathbf{v}_1), T(\mathbf{v}_2), \ldots, T(\mathbf{v}_n)$ are linearly independent elements of W. This proves (i).

Next suppose that $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ are elements of V that span V. Let $\mathbf{w}$ be any element of W. Then $\mathbf{w} = T(\mathbf{u})$ for some element $\mathbf{u}$ of V, since the isomorphism $T: V \to W$ is surjective. There exist scalars $c_1, c_2, \ldots, c_n$ for which $\mathbf{u} = \sum_{j=1}^{n} c_j \mathbf{v}_j$, since the elements $\mathbf{v}_j$ span V, and then

$$\mathbf{w} = T(\mathbf{u}) = T\left(\sum_{j=1}^{n} c_j \mathbf{v}_j\right) = \sum_{j=1}^{n} c_j T(\mathbf{v}_j).$$

Thus $T(\mathbf{v}_1), T(\mathbf{v}_2), \ldots, T(\mathbf{v}_n)$ span the vector space W. This proves (ii).

Finally we note that (iii) is an immediate consequence of (i) and (ii). ∎

Theorem 9.25. *Two finite-dimensional vector spaces over a given field are isomorphic if and only if they have the same dimension.*

Proof. The dimension of a vector space is the number of elements in any basis of that vector space. It follows immediately from Lemma 9.24 that isomorphic vector spaces have the same dimension.

Suppose that V and W are vector spaces over a given field that have the same dimension. We must show that V and W are isomorphic. This result clearly holds when the dimension of both spaces is zero. Suppose then that the dimension of these spaces is non-zero. Now there exists a basis $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ of V and a basis $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n$ of W where both bases have the same number of elements. Moreover, given any element $\mathbf{u}$ of V, there exist uniquely determined scalars $x_1, x_2, \ldots, x_n$ such that $\mathbf{u} = \sum_{j=1}^{n} x_j \mathbf{v}_j$ (Theorem 9.7). It follows that there exists a well-defined function $T: V \to W$ characterized by the property that $T\left(\sum_{j=1}^{n} x_j \mathbf{v}_j\right) = \sum_{j=1}^{n} x_j \mathbf{w}_j$ for all scalars $x_1, x_2, \ldots, x_n$. One

can readily check that the function T is a linear transformation. It is also invertible: the inverse T^{-1} of T satisfies $T^{-1}\left(\sum_{j=1}^{n} x_j \mathbf{w}_j\right) = \sum_{j=1}^{n} x_j \mathbf{v}_j$ for all scalars $x_1, x_2, \ldots, x_n$. Therefore $T: V \to W$ is an isomorphism, and thus the vector spaces V and W are isomorphic, as required. ∎

Let K be a field, and, for each positive integer n, let K^n be the set of all ordered n-tuples $(x_1, x_2, \ldots, x_n)$ of elements of K. Then K^n is a vector space over the field K, where

$$(x_1, x_2, \ldots, x_n) + (y_1, y_2, \ldots, y_n) = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n)$$

and

$$c(x_1, x_2, \ldots, x_n) = (cx_1, cx_2, \ldots, cx_n)$$

for all elements $(x_1, x_2, \ldots, x_n)$ and $(y_1, y_2, \ldots, y_n)$ of K^n and for all elements c of K. Now K^n is an n-dimensional vector space over K, and the ordered n-tuples in which one component has the value 1 and the remaining components are zero constitute a basis of K^n. Theorem 9.25 ensures that any vector space of dimension n over the field K is isomorphic to the vector space K^n. Indeed if $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ is a basis for a vector space V of dimension n over the field K, then the function sending an element $(x_1, x_2, \ldots, x_n)$ of K^n to $\sum_{j=1}^{n} x_j \mathbf{v}_j$ is an isomorphism from K^n to V.

Any real vector space of dimension n is isomorphic to the space $\mathbb{R}^n$. In particular, any 3-dimensional real vector space is isomorphic to the vector space consisting of all vectors in 3-dimensional Euclidean space.

Problems

1. Let V be a vector space over a field K. Using the vector space axioms, the definition of subtraction in a vector space etc., prove formally that the following identities are satisfied for all elements $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ of V and for all elements c and d of V:

   (i) $\mathbf{u} - (\mathbf{v} - \mathbf{w}) = (\mathbf{u} - \mathbf{v}) + \mathbf{w}$;

   (ii) $c(\mathbf{v} - \mathbf{w}) = c\mathbf{v} - c\mathbf{w}$;

   (iii) $(c - d)\mathbf{v} = c\mathbf{v} - d\mathbf{v}$.

2. Let $C([0, 1], \mathbb{R})$ be the real vector space consisting of all continuous real-valued functions on the interval $[0, 1]$, where the operations of addition of functions and of multiplication of functions by real numbers are defined in the usual fashion. Is the subset of $C([0, 1], \mathbb{R})$ consisting of those continuous functions $f: [0, 1] \to \mathbb{R}$ that satisfy $0 \leq \int_0^1 f(x)\, dx \leq 1$ a subspace of $C([0, 1], \mathbb{R})$?

3. Show that $(1, 0, 0)$, $(1, 2, 0)$, $(1, 2, 3)$ is a basis of the real vector space $\mathbb{R}^3$.

4. For each non-negative integer n, let V_n be the real vector space consisting of all polynomials with real coefficients whose degree does not exceed n.

   (i) Show that $24, 24x, 12x^2, 4x^3, x^4$ is a basis of V_4.

   (ii) Suppose that $n \geq 1$. Is the function from V_n to itself that sends a polynomial $p(x)$ to $p(x) + x$ a linear transformation?

   (iii) Is the function from V_n to V_{n+1} that sends a polynomial $p(x)$ to $xp(x)$ a linear transformation?

(iv) Let $S: V_4 \to V_4$ be the linear transformation from V_4 to itself that sends a polynomial $p(x)$ to its derivative $p'(x)$. Write down the matrix that represents S with respect to the basis $24, 24x, 12x^2, 4x^3, x^4$ of V_4.

(v) Let $T: V_4 \to V_3$ be the linear transformation from V_4 to V_3 that sends a polynomial $p(x)$ to $p''(x) - 3p'(x)$. Calculate the matrix that represents T with respect to the basis $1, x, x^2, x^3, x^4$ of V_4 and the basis $1, x, x^2, x^3$ of V_3.

5. Let $T: \mathbb{R}^3 \to \mathbb{R}^2$ be the linear transformation given by $T(u_1, u_2, u_3) = (u_1 + u_2, u_1 - u_2)$. Find a basis $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ of $\mathbb{R}^3$ such that $T(\mathbf{v}_1) = (1, 0)$, $T(\mathbf{v}_2) = (0, 1)$ and $T(\mathbf{v}_3) = (0, 0)$.

6. Let U, V and W be vector spaces over some given field. Suppose that the vector spaces U and V are isomorphic and that the vector spaces V and W are isomorphic. Explain why the vector spaces U and W are then isomorphic.

7. Let V be a vector space of dimension n over a field K. Suppose that the field K is finite and has q elements. Is the vector space V then a finite set, and, if so, how many elements does it have?

8. Let V be a vector space over a field K, and let U be a subspace of V. A *coset* of U in V is a subset of V that is of the form $U + \mathbf{v}$, where $U + \mathbf{v} = \{\mathbf{u} + \mathbf{v} : \mathbf{u} \in U\}$. Show that the set V/U of all cosets of U in V is a well-defined vector space over K, where $(U + \mathbf{v}) + (U + \mathbf{w}) = U + \mathbf{v} + \mathbf{w}$ and $c(U + \mathbf{v}) = U + c\mathbf{v}$ for all elements $\mathbf{v}$ and $\mathbf{w}$ of V and for all elements c of K.

   Now let V and W be vector spaces over K, and let $T: V \to W$ be a linear transformation. Show that the function that sends a coset $\ker T + \mathbf{v}$ of the kernel $\ker T$ to $T(\mathbf{v})$ is well-defined and is an isomorphism from $V/\ker T$ to $T(V)$.