Chapter 6

Vector Spaces

In this chapter we introduce vector spaces in full generality. The reader will notice some similarity with the discussion of the space $\mathbb{R}^n$ in Chapter 5. In fact much of the present material has been developed in that context, and there is some repetition. However, Chapter 6 deals with the notion of an *abstract* vector space, a concept that will be new to most readers. It turns out that there are many systems in which a natural addition and scalar multiplication are defined and satisfy the usual rules familiar from $\mathbb{R}^n$. The study of abstract vector spaces is a way to deal with all these examples simultaneously. The new aspect is that we are dealing with an abstract system in which *all we know* about the vectors is that they are objects that can be added and multiplied by a scalar and satisfy rules familiar from $\mathbb{R}^n$.

The novel thing is the *abstraction*. Getting used to this new conceptual level is facilitated by the work done in Chapter 5: First, the vector manipulations are familiar, giving the reader more time to become accustomed to the abstract setting; and, second, the mental images developed in the concrete setting of $\mathbb{R}^n$ serve as an aid to doing many of the exercises in Chapter 6.

The concept of a vector space was first introduced in 1844 by the German mathematician Hermann Grassmann (1809-1877), but his work did not receive the attention it deserved. It was not until 1888 that the Italian mathematician Guiseppe Peano (1858-1932) clarified Grassmann's work in his book *Calcolo Geometrico* and gave the vector space axioms in their present form. Vector spaces became established with the work of the Polish mathematician Stephan Banach (1892-1945), and the idea was finally accepted in 1918 when Hermann Weyl (1885-1955) used it in his widely read book *Raum-Zeit-Materie* ("Space-Time-Matter"), an introduction to the general theory of relativity.

6.1 Examples and Basic Properties

Many mathematical entities have the property that they can be added and multiplied by a number. Numbers themselves have this property, as do $m \times n$ matrices: The sum of two such matrices is again $m \times n$ as is any scalar multiple of such a matrix. Polynomials are another familiar example, as are the geometric vectors in Chapter 4. It turns out that there are many other types of mathematical objects that can be added and multiplied by a scalar, and the general study of such systems is introduced in this chapter. Remarkably, much of what we could say in Chapter 5 about the dimension of subspaces in $\mathbb{R}^n$ can be formulated in this generality.

Definition 6.1 Vector Spaces

*A **vector space** consists of a nonempty set V of objects (called **vectors**) that can be added, that can be multiplied by a real number (called a **scalar** in this context), and for which certain axioms hold.*[1] *If $\mathbf{v}$ and $\mathbf{w}$ are two vectors in V, their sum is expressed as $\mathbf{v}+\mathbf{w}$, and the scalar product of $\mathbf{v}$ by a real number a is denoted as $a\mathbf{v}$. These operations are called **vector addition** and **scalar multiplication**, respectively, and the following axioms are assumed to hold.*

Axioms for vector addition

A1. *If $\mathbf{u}$ and $\mathbf{v}$ are in V, then $\mathbf{u}+\mathbf{v}$ is in V.*

A2. $\mathbf{u}+\mathbf{v}=\mathbf{v}+\mathbf{u}$ *for all $\mathbf{u}$ and $\mathbf{v}$ in V.*

A3. $\mathbf{u}+(\mathbf{v}+\mathbf{w})=(\mathbf{u}+\mathbf{v})+\mathbf{w}$ *for all $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ in V.*

A4. *An element $\mathbf{0}$ in V exists such that $\mathbf{v}+\mathbf{0}=\mathbf{v}=\mathbf{0}+\mathbf{v}$ for every $\mathbf{v}$ in V.*

A5. *For each $\mathbf{v}$ in V, an element $-\mathbf{v}$ in V exists such that $-\mathbf{v}+\mathbf{v}=\mathbf{0}$ and $\mathbf{v}+(-\mathbf{v})=\mathbf{0}$.*

Axioms for scalar multiplication

S1. *If $\mathbf{v}$ is in V, then $a\mathbf{v}$ is in V for all a in $\mathbb{R}$.*

S2. $a(\mathbf{v}+\mathbf{w})=a\mathbf{v}+a\mathbf{w}$ *for all $\mathbf{v}$ and $\mathbf{w}$ in V and all a in $\mathbb{R}$.*

S3. $(a+b)\mathbf{v}=a\mathbf{v}+b\mathbf{v}$ *for all $\mathbf{v}$ in V and all a and b in $\mathbb{R}$.*

S4. $a(b\mathbf{v})=(ab)\mathbf{v}$ *for all $\mathbf{v}$ in V and all a and b in $\mathbb{R}$.*

S5. $1\mathbf{v}=\mathbf{v}$ *for all $\mathbf{v}$ in V.*

The content of axioms A1 and S1 is described by saying that V is **closed** under vector addition and scalar multiplication. The element $\mathbf{0}$ in axiom A4 is called the **zero vector**, and the vector $-\mathbf{v}$ in axiom A5 is called the **negative** of $\mathbf{v}$.

The rules of matrix arithmetic, when applied to $\mathbb{R}^n$, give

Example 6.1.1

$\mathbb{R}^n$ is a vector space using matrix addition and scalar multiplication.[2]

It is important to realize that, in a general vector space, the vectors need not be n-tuples as in $\mathbb{R}^n$. They can be any kind of objects at all as long as the addition and scalar multiplication are defined and the axioms are satisfied. The following examples illustrate the diversity of the concept.

The space $\mathbb{R}^n$ consists of special types of matrices. More generally, let $\mathbf{M}_{mn}$ denote the set of all $m\times n$ matrices with real entries. Then Theorem 2.1.1 gives:

---

[1] The scalars will usually be real numbers, but they could be complex numbers, or elements of an algebraic system called a field. Another example is the field $\mathbb{Q}$ of rational numbers. We will look briefly at finite fields in Section 8.8.

[2] We will usually write the vectors in $\mathbb{R}^n$ as n-tuples. However, if it is convenient, we will sometimes denote them as rows or columns.

Example 6.1.2

The set $\mathbf{M}_{mn}$ of all $m \times n$ matrices is a vector space using matrix addition and scalar multiplication. The zero element in this vector space is the zero matrix of size $m \times n$, and the vector space negative of a matrix (required by axiom A5) is the usual matrix negative discussed in Section 2.1. Note that $\mathbf{M}_{mn}$ is just $\mathbb{R}^{mn}$ in different notation.

In Chapter 5 we identified many important subspaces of $\mathbb{R}^n$ such as $\operatorname{im} A$ and $\operatorname{null} A$ for a matrix A. These are all vector spaces.

Example 6.1.3

Show that every subspace of $\mathbb{R}^n$ is a vector space in its own right using the addition and scalar multiplication of $\mathbb{R}^n$.

Solution. Axioms A1 and S1 are two of the defining conditions for a subspace U of $\mathbb{R}^n$ (see Section 5.1). The other eight axioms for a vector space are inherited from $\mathbb{R}^n$. For example, if $\mathbf{x}$ and $\mathbf{y}$ are in U and a is a scalar, then $a(\mathbf{x}+\mathbf{y}) = a\mathbf{x}+a\mathbf{y}$ because $\mathbf{x}$ and $\mathbf{y}$ are in $\mathbb{R}^n$. This shows that axiom S2 holds for U; similarly, the other axioms also hold for U.

Example 6.1.4

Let V denote the set of all ordered pairs (x, y) and define addition in V as in $\mathbb{R}^2$. However, define a new scalar multiplication in V by

$$a(x, y) = (ay, ax)$$

Determine if V is a vector space with these operations.

Solution. Axioms A1 to A5 are valid for V because they hold for matrices. Also $a(x, y) = (ay, ax)$ is again in V, so axiom S1 holds. To verify axiom S2, let $\mathbf{v} = (x, y)$ and $\mathbf{w} = (x_1, y_1)$ be typical elements in V and compute

$$a(\mathbf{v}+\mathbf{w}) = a(x+x_1, y+y_1) = (a(y+y_1), a(x+x_1))$$
$$a\mathbf{v}+a\mathbf{w} = (ay, ax)+(ay_1, ax_1) = (ay+ay_1, ax+ax_1)$$

Because these are equal, axiom S2 holds. Similarly, the reader can verify that axiom S3 holds. However, axiom S4 fails because

$$a(b(x, y)) = a(by, bx) = (abx, aby)$$

need not equal $ab(x, y) = (aby, abx)$. Hence, V is *not* a vector space. (In fact, axiom S5 also fails.)

Sets of polynomials provide another important source of examples of vector spaces, so we review some basic facts. A **polynomial** in an indeterminate x is an expression

$$p(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$$

where $a_0, a_1, a_2, \dots, a_n$ are real numbers called the **coefficients** of the polynomial. If all the coefficients are zero, the polynomial is called the **zero polynomial** and is denoted simply as 0. If $p(x) \neq 0$, the highest power of x with a nonzero coefficient is called the **degree** of $p(x)$ denoted as $\deg p(x)$. The coefficient itself is called the **leading coefficient** of $p(x)$. Hence $\deg(3+5x) = 1$, $\deg(1+x+x^2) = 2$, and $\deg(4) = 0$. (The degree of the zero polynomial is not defined.)

Let **P** denote the set of all polynomials and suppose that

$$\begin{aligned} p(x) &= a_0 + a_1x + a_2x^2 + \cdots \\ q(x) &= b_0 + b_1x + b_2x^2 + \cdots \end{aligned}$$

are two polynomials in **P** (possibly of different degrees). Then $p(x)$ and $q(x)$ are called **equal** [written $p(x) = q(x)$] if and only if all the corresponding coefficients are equal—that is, $a_0 = b_0$, $a_1 = b_1$, $a_2 = b_2$, and so on. In particular, $a_0 + a_1x + a_2x^2 + \cdots = 0$ means that $a_0 = 0$, $a_1 = 0$, $a_2 = 0$, $\dots$, and this is the reason for calling x an **indeterminate**. The set **P** has an addition and scalar multiplication defined on it as follows: if $p(x)$ and $q(x)$ are as before and a is a real number,

$$\begin{aligned} p(x) + q(x) &= (a_0 + b_0) + (a_1 + b_1)x + (a_2 + b_2)x^2 + \cdots \\ ap(x) &= aa_0 + (aa_1)x + (aa_2)x^2 + \cdots \end{aligned}$$

Evidently, these are again polynomials, so **P** is closed under these operations, called **pointwise** addition and scalar multiplication. The other vector space axioms are easily verified, and we have

Example 6.1.5

The set **P** of all polynomials is a vector space with the foregoing addition and scalar multiplication. The zero vector is the zero polynomial, and the negative of a polynomial $p(x) = a_0 + a_1x + a_2x^2 + \dots$ is the polynomial $-p(x) = -a_0 - a_1x - a_2x^2 - \dots$ obtained by negating all the coefficients.

There is another vector space of polynomials that will be referred to later.

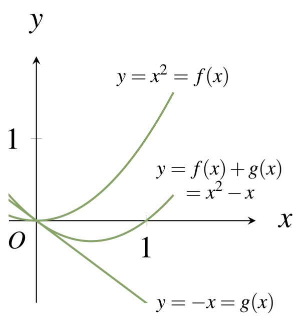
Example 6.1.6

Given $n \geq 1$, let $\mathbf{P}_n$ denote the set of all polynomials of degree at most n, together with the zero polynomial. That is

$$\mathbf{P}_n = \{a_0 + a_1x + a_2x^2 + \cdots + a_nx^n \mid a_0,\ a_1,\ a_2,\ \dots,\ a_n \text{ in } \mathbb{R}\}.$$

Then $\mathbf{P}_n$ is a vector space. Indeed, sums and scalar multiples of polynomials in $\mathbf{P}_n$ are again in $\mathbf{P}_n$, and the other vector space axioms are inherited from **P**. In particular, the zero vector and the negative of a polynomial in $\mathbf{P}_n$ are the same as those in **P**.

If a and b are real numbers and $a < b$, the **interval** $[a,\ b]$ is defined to be the set of all real numbers x such that $a \leq x \leq b$. A (real-valued) **function** f on $[a,\ b]$ is a rule that associates to every number x in $[a,\ b]$ a real number denoted $f(x)$. The rule is frequently specified by giving a formula for $f(x)$ in terms of x. For example, $f(x) = 2^x$, $f(x) = \sin x$, and $f(x) = x^2 + 1$ are familiar functions. In fact, every polynomial $p(x)$ can be regarded as the formula for a function p.



The set of all functions on $[a, b]$ is denoted $\mathbf{F}[a, b]$. Two functions f and g in $\mathbf{F}[a, b]$ are **equal** if $f(x) = g(x)$ for every x in $[a, b]$, and we describe this by saying that f and g have the **same action**. Note that two polynomials are equal in $\mathbf{P}$ (defined prior to Example 6.1.5) if and only if they are equal as functions.

If f and g are two functions in $\mathbf{F}[a, b]$, and if r is a real number, define the sum $f + g$ and the scalar product rf by

$$\begin{aligned} (f+g)(x) &= f(x) + g(x) \quad \text{for each } x \text{ in } [a, b] \\ (rf)(x) &= rf(x) \quad \text{for each } x \text{ in } [a, b] \end{aligned}$$

In other words, the action of $f + g$ upon x is to associate x with the number $f(x) + g(x)$, and rf associates x with $rf(x)$. The sum of $f(x) = x^2$ and $g(x) = -x$ is shown in the diagram. These operations on $\mathbf{F}[a, b]$ are called **pointwise addition and scalar multiplication** of functions and they are the usual operations familiar from elementary algebra and calculus.

Example 6.1.7

The set $\mathbf{F}[a, b]$ of all functions on the interval $[a, b]$ is a vector space using pointwise addition and scalar multiplication. The zero function (in axiom A4), denoted 0, is the constant function defined by

$$0(x) = 0 \quad \text{for each } x \text{ in } [a, b]$$

The negative of a function f is denoted $-f$ and has action defined by

$$(-f)(x) = -f(x) \quad \text{for each } x \text{ in } [a, b]$$

Axioms A1 and S1 are clearly satisfied because, if f and g are functions on $[a, b]$, then $f + g$ and rf are again such functions. The verification of the remaining axioms is left as Exercise 6.1.14.

Other examples of vector spaces will appear later, but these are sufficiently varied to indicate the scope of the concept and to illustrate the properties of vector spaces to be discussed. With such a variety of examples, it may come as a surprise that a well-developed *theory* of vector spaces exists. That is, many properties can be shown to hold for *all* vector spaces and hence hold in every example. Such properties are called *theorems* and can be deduced from the axioms. Here is an important example.

Theorem 6.1.1: Cancellation

Let $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ be vectors in a vector space V. If $\mathbf{v} + \mathbf{u} = \mathbf{v} + \mathbf{w}$, then $\mathbf{u} = \mathbf{w}$.

Proof. We are given $\mathbf{v} + \mathbf{u} = \mathbf{v} + \mathbf{w}$. If these were numbers instead of vectors, we would simply subtract $\mathbf{v}$ from both sides of the equation to obtain $\mathbf{u} = \mathbf{w}$. This can be accomplished with vectors by adding $-\mathbf{v}$ to both sides of the equation. The steps (using only the axioms) are as follows:

$$\begin{aligned} \mathbf{v} + \mathbf{u} &= \mathbf{v} + \mathbf{w} \\ -\mathbf{v} + (\mathbf{v} + \mathbf{u}) &= -\mathbf{v} + (\mathbf{v} + \mathbf{w}) && \text{(axiom A5)} \\ (-\mathbf{v} + \mathbf{v}) + \mathbf{u} &= (-\mathbf{v} + \mathbf{v}) + \mathbf{w} && \text{(axiom A3)} \end{aligned}$$

$$\mathbf{0}+\mathbf{u}=\mathbf{0}+\mathbf{w} \quad \text{(axiom A5)}$$
$$\mathbf{u}=\mathbf{w} \quad \text{(axiom A4)}$$

This is the desired conclusion.[3] □

As with many good mathematical theorems, the technique of the proof of Theorem 6.1.1 is at least as important as the theorem itself. The idea was to mimic the well-known process of numerical subtraction in a vector space V as follows: To subtract a vector $\mathbf{v}$ from both sides of a vector equation, we added $-\mathbf{v}$ to both sides. With this in mind, we define **difference** $\mathbf{u}-\mathbf{v}$ of two vectors in V as

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

We shall say that this vector is the result of having **subtracted** $\mathbf{v}$ from $\mathbf{u}$ and, as in arithmetic, this operation has the property given in Theorem 6.1.2.

Theorem 6.1.2

If $\mathbf{u}$ and $\mathbf{v}$ are vectors in a vector space V, the equation

$$\mathbf{x}+\mathbf{v}=\mathbf{u}$$

has one and only one solution $\mathbf{x}$ in V given by

$$\mathbf{x}=\mathbf{u}-\mathbf{v}$$

Proof. The difference $\mathbf{x}=\mathbf{u}-\mathbf{v}$ is indeed a solution to the equation because (using several axioms)

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

To see that this is the only solution, suppose $\mathbf{x}_1$ is another solution so that $\mathbf{x}_1+\mathbf{v}=\mathbf{u}$. Then $\mathbf{x}+\mathbf{v}=\mathbf{x}_1+\mathbf{v}$ (they both equal $\mathbf{u}$), so $\mathbf{x}=\mathbf{x}_1$ by cancellation. □

Similarly, cancellation shows that there is only one zero vector in any vector space and only one negative of each vector (Exercises 6.1.10 and 6.1.11). Hence we speak of *the* zero vector and *the* negative of a vector.

The next theorem derives some basic properties of scalar multiplication that hold in every vector space, and will be used extensively.

Theorem 6.1.3

Let $\mathbf{v}$ denote a vector in a vector space V and let a denote a real number.

1. $0\mathbf{v}=\mathbf{0}$.
2. $a\mathbf{0}=\mathbf{0}$.
3. *If $a\mathbf{v}=\mathbf{0}$, then either $a=0$ or $\mathbf{v}=\mathbf{0}$.*
4. $(-1)\mathbf{v}=-\mathbf{v}$.

[3] Observe that none of the scalar multiplication axioms are needed here.

5. $(-a)\mathbf{v} = -(a\mathbf{v}) = a(-\mathbf{v})$.

Proof.

1. Observe that $0\mathbf{v} + 0\mathbf{v} = (0+0)\mathbf{v} = 0\mathbf{v} = 0\mathbf{v} + \mathbf{0}$ where the first equality is by axiom S3. It follows that $0\mathbf{v} = \mathbf{0}$ by cancellation.

2. The proof is similar to that of (1), and is left as Exercise 6.1.12(a).

3. Assume that $a\mathbf{v} = \mathbf{0}$. If $a = 0$, there is nothing to prove; if $a \neq 0$, we must show that $\mathbf{v} = \mathbf{0}$. But $a \neq 0$ means we can scalar-multiply the equation $a\mathbf{v} = \mathbf{0}$ by the scalar $\frac{1}{a}$. The result (using (2) and Axioms S5 and S4) is
$$\mathbf{v} = 1\mathbf{v} = \left(\frac{1}{a}a\right)\mathbf{v} = \frac{1}{a}(a\mathbf{v}) = \frac{1}{a}\mathbf{0} = \mathbf{0}$$

4. We have $-\mathbf{v} + \mathbf{v} = \mathbf{0}$ by axiom A5. On the other hand,
$$(-1)\mathbf{v} + \mathbf{v} = (-1)\mathbf{v} + 1\mathbf{v} = (-1+1)\mathbf{v} = 0\mathbf{v} = \mathbf{0}$$
using (1) and axioms S5 and S3. Hence $(-1)\mathbf{v} + \mathbf{v} = -\mathbf{v} + \mathbf{v}$ (because both are equal to $\mathbf{0}$), so $(-1)\mathbf{v} = -\mathbf{v}$ by cancellation.

5. The proof is left as Exercise 6.1.12.[4] □

The properties in Theorem 6.1.3 are familiar for matrices; the point here is that they hold in *every* vector space. It is hard to exaggerate the importance of this observation.

Axiom A3 ensures that the sum $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$ is the same however it is formed, and we write it simply as $\mathbf{u} + \mathbf{v} + \mathbf{w}$. Similarly, there are different ways to form any sum $\mathbf{v}_1 + \mathbf{v}_2 + \cdots + \mathbf{v}_n$, and Axiom A3 guarantees that they are all equal. Moreover, Axiom A2 shows that the order in which the vectors are written does not matter (for example: $\mathbf{u} + \mathbf{v} + \mathbf{w} + \mathbf{z} = \mathbf{z} + \mathbf{u} + \mathbf{w} + \mathbf{v}$).

Similarly, Axioms S2 and S3 extend. For example
$$a(\mathbf{u} + \mathbf{v} + \mathbf{w}) = a[\mathbf{u} + (\mathbf{v} + \mathbf{w})] = a\mathbf{u} + a(\mathbf{v} + \mathbf{w}) = a\mathbf{u} + a\mathbf{v} + a\mathbf{w}$$
for all a, $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$. Similarly $(a+b+c)\mathbf{v} = a\mathbf{v} + b\mathbf{v} + c\mathbf{v}$ hold for all values of a, b, c, and $\mathbf{v}$ (verify). More generally,
$$\begin{aligned} a(\mathbf{v}_1 + \mathbf{v}_2 + \cdots + \mathbf{v}_n) &= a\mathbf{v}_1 + a\mathbf{v}_2 + \cdots + a\mathbf{v}_n \\ (a_1 + a_2 + \cdots + a_n)\mathbf{v} &= a_1\mathbf{v} + a_2\mathbf{v} + \cdots + a_n\mathbf{v} \end{aligned}$$
hold for all $n \geq 1$, all numbers $a, a_1, \dots, a_n$, and all vectors, $\mathbf{v}, \mathbf{v}_1, \dots, \mathbf{v}_n$. The verifications are by induction and are left to the reader (Exercise 6.1.13). These facts—together with the axioms, Theorem 6.1.3, and the definition of subtraction—enable us to simplify expressions involving sums of scalar multiples of vectors by collecting like terms, expanding, and taking out common factors. This has been discussed for the vector space of matrices in Section 2.1 (and for geometric vectors in Section 4.1); the manipulations in an arbitrary vector space are carried out in the same way. Here is an illustration.

Example 6.1.8

If **u**, **v**, and **w** are vectors in a vector space V, simplify the expression

$$2(\mathbf{u}+3\mathbf{w})-3(2\mathbf{w}-\mathbf{v})-3[2(2\mathbf{u}+\mathbf{v}-4\mathbf{w})-4(\mathbf{u}-2\mathbf{w})]$$

Solution. The reduction proceeds as though **u**, **v**, and **w** were matrices or variables.

$$\begin{aligned}
&2(\mathbf{u}+3\mathbf{w})-3(2\mathbf{w}-\mathbf{v})-3[2(2\mathbf{u}+\mathbf{v}-4\mathbf{w})-4(\mathbf{u}-2\mathbf{w})] \\
&=2\mathbf{u}+6\mathbf{w}-6\mathbf{w}+3\mathbf{v}-3[4\mathbf{u}+2\mathbf{v}-8\mathbf{w}-4\mathbf{u}+8\mathbf{w}] \\
&=2\mathbf{u}+3\mathbf{v}-3[2\mathbf{v}] \\
&=2\mathbf{u}+3\mathbf{v}-6\mathbf{v} \\
&=2\mathbf{u}-3\mathbf{v}
\end{aligned}$$

Condition (2) in Theorem 6.1.3 points to another example of a vector space.

Example 6.1.9

A set $\{\mathbf{0}\}$ with one element becomes a vector space if we define

$$\mathbf{0}+\mathbf{0}=\mathbf{0} \quad \text{and} \quad a\mathbf{0}=\mathbf{0} \text{ for all scalars } a.$$

The resulting space is called the **zero vector space** and is denoted $\{\mathbf{0}\}$.

The vector space axioms are easily verified for $\{\mathbf{0}\}$. In any vector space V, Theorem 6.1.3 shows that the zero subspace (consisting of the zero vector of V alone) is a copy of the zero vector space.

Exercises for 6.1

Exercise 6.1.1 Let V denote the set of ordered triples (x, y, z) and define addition in V as in $\mathbb{R}^3$. For each of the following definitions of scalar multiplication, decide whether V is a vector space.

a. $a(x, y, z)=(ax, y, az)$

b. $a(x, y, z)=(ax, 0, az)$

c. $a(x, y, z)=(0, 0, 0)$

d. $a(x, y, z)=(2ax, 2ay, 2az)$

Exercise 6.1.2 Are the following sets vector spaces with the indicated operations? If not, why not?

a. The set V of nonnegative real numbers; ordinary addition and scalar multiplication.

b. The set V of all polynomials of degree ≥ 3, together with 0; operations of **P**.

c. The set of all polynomials of degree ≤ 3; operations of **P**.

d. The set $\{1, x, x^2, \dots\}$; operations of **P**.

e. The set V of all 2×2 matrices of the form $\left[\begin{array}{cc} a & b \\ 0 & c \end{array}\right]$; operations of $\mathbf{M}_{22}$.

f. The set V of 2×2 matrices with equal column sums; operations of $\mathbf{M}_{22}$.

g. The set V of 2×2 matrices with zero determinant; usual matrix operations.

h. The set V of real numbers; usual operations.

i. The set V of complex numbers; usual addition and multiplication by a real number.

j. The set V of all ordered pairs (x, y) with the addition of $\mathbb{R}^2$, but using scalar multiplication $a(x, y) = (ax, -ay)$.

k. The set V of all ordered pairs (x, y) with the addition of $\mathbb{R}^2$, but using scalar multiplication $a(x, y) = (x, y)$ for all a in $\mathbb{R}$.

l. The set V of all functions $f : \mathbb{R} \to \mathbb{R}$ with pointwise addition, but scalar multiplication defined by $(af)(x) = f(ax)$.

m. The set V of all 2×2 matrices whose entries sum to 0; operations of $\mathbf{M}_{22}$.

n. The set V of all 2×2 matrices with the addition of $\mathbf{M}_{22}$ but scalar multiplication $*$ defined by $a * X = aX^T$.

Exercise 6.1.3 Let V be the set of positive real numbers with vector addition being ordinary multiplication, and scalar multiplication being $a \cdot v = v^a$. Show that V is a vector space.

Exercise 6.1.4 If V is the set of ordered pairs (x, y) of real numbers, show that it is a vector space with addition $(x, y) + (x_1, y_1) = (x + x_1, y + y_1 + 1)$ and scalar multiplication $a(x, y) = (ax, ay + a - 1)$. What is the zero vector in V?

Exercise 6.1.5 Find $\mathbf{x}$ and $\mathbf{y}$ (in terms of $\mathbf{u}$ and $\mathbf{v}$) such that:

a. $\begin{aligned} 2\mathbf{x} + \mathbf{y} &= \mathbf{u} \\ 5\mathbf{x} + 3\mathbf{y} &= \mathbf{v} \end{aligned}$

b. $\begin{aligned} 3\mathbf{x} - 2\mathbf{y} &= \mathbf{u} \\ 4\mathbf{x} - 5\mathbf{y} &= \mathbf{v} \end{aligned}$

Exercise 6.1.6 In each case show that the condition $a\mathbf{u} + b\mathbf{v} + c\mathbf{w} = \mathbf{0}$ in V implies that $a = b = c = 0$.

a. $V = \mathbb{R}^4$; $\mathbf{u} = (2, 1, 0, 2)$, $\mathbf{v} = (1, 1, -1, 0)$, $\mathbf{w} = (0, 1, 2, 1)$

b. $V = \mathbf{M}_{22}$; $\mathbf{u} = \left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array} \right]$, $\mathbf{v} = \left[\begin{array}{cc} 0 & 1 \\ 1 & 0 \end{array} \right]$, $\mathbf{w} = \left[\begin{array}{cc} 1 & 1 \\ 1 & -1 \end{array} \right]$

c. $V = \mathbf{P}$; $\mathbf{u} = x^3 + x$, $\mathbf{v} = x^2 + 1$, $\mathbf{w} = x^3 - x^2 + x + 1$

d. $V = \mathbf{F}[0, \pi]$; $\mathbf{u} = \sin x$, $\mathbf{v} = \cos x$, $\mathbf{w} = 1$—the constant function

Exercise 6.1.7 Simplify each of the following.

a. $3[2(\mathbf{u} - 2\mathbf{v} - \mathbf{w}) + 3(\mathbf{w} - \mathbf{v})] - 7(\mathbf{u} - 3\mathbf{v} - \mathbf{w})$

b. $4(3\mathbf{u} - \mathbf{v} + \mathbf{w}) - 2[(3\mathbf{u} - 2\mathbf{v}) - 3(\mathbf{v} - \mathbf{w})]$ $+ 6(\mathbf{w} - \mathbf{u} - \mathbf{v})$

Exercise 6.1.8 Show that $\mathbf{x} = \mathbf{v}$ is the only solution to the equation $\mathbf{x} + \mathbf{x} = 2\mathbf{v}$ in a vector space V. Cite all axioms used.

Exercise 6.1.9 Show that $-\mathbf{0} = \mathbf{0}$ in any vector space. Cite all axioms used.

Exercise 6.1.10 Show that the zero vector $\mathbf{0}$ is uniquely determined by the property in axiom A4.

Exercise 6.1.11 Given a vector $\mathbf{v}$, show that its negative $-\mathbf{v}$ is uniquely determined by the property in axiom A5.

Exercise 6.1.12

a. Prove (2) of Theorem 6.1.3. [*Hint*: Axiom S2.]

b. Prove that $(-a)\mathbf{v} = -(a\mathbf{v})$ in Theorem 6.1.3 by first computing $(-a)\mathbf{v} + a\mathbf{v}$. Then do it using (4) of Theorem 6.1.3 and axiom S4.

c. Prove that $a(-\mathbf{v}) = -(a\mathbf{v})$ in Theorem 6.1.3 in two ways, as in part (b).

Exercise 6.1.13 Let $\mathbf{v}, \mathbf{v}_1, \dots, \mathbf{v}_n$ denote vectors in a vector space V and let $a, a_1, \dots, a_n$ denote numbers. Use induction on n to prove each of the following.

a. $a(\mathbf{v}_1 + \mathbf{v}_2 + \cdots + \mathbf{v}_n) = a\mathbf{v}_1 + a\mathbf{v}_2 + \cdots + a\mathbf{v}_n$

b. $(a_1 + a_2 + \cdots + a_n)\mathbf{v} = a_1\mathbf{v} + a_2\mathbf{v} + \cdots + a_n\mathbf{v}$

Exercise 6.1.14 Verify axioms A2—A5 and S2—S5 for the space $\mathbf{F}[a, b]$ of functions on $[a, b]$ (Example 6.1.7).

Exercise 6.1.15 Prove each of the following for vectors $\mathbf{u}$ and $\mathbf{v}$ and scalars a and b.

a. If $a\mathbf{v} = \mathbf{0}$, then $a = 0$ or $\mathbf{v} = \mathbf{0}$.

b. If $a\mathbf{v} = b\mathbf{v}$ and $\mathbf{v} \neq \mathbf{0}$, then $a = b$.

c. If $a\mathbf{v} = a\mathbf{w}$ and $a \neq 0$, then $\mathbf{v} = \mathbf{w}$.

Exercise 6.1.16 By calculating $(1+1)(\mathbf{v}+\mathbf{w})$ in two ways (using axioms S2 and S3), show that axiom A2 follows from the other axioms.

Exercise 6.1.17 Let V be a vector space, and define V^n to be the set of all n-tuples $(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n)$ of n vectors $\mathbf{v}_i$, each belonging to V. Define addition and scalar multiplication in V^n as follows:

$$(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n) + (\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n) = (\mathbf{u}_1+\mathbf{v}_1, \mathbf{u}_2+\mathbf{v}_2, \dots, \mathbf{u}_n+\mathbf{v}_n)$$
$$a(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n) = (a\mathbf{v}_1, a\mathbf{v}_2, \dots, a\mathbf{v}_n)$$

Show that V^n is a vector space.

Exercise 6.1.18 Let V^n be the vector space of n-tuples from the preceding exercise, written as columns. If A is an $m \times n$ matrix, and X is in V^n, define AX in V^m by matrix multiplication. More precisely, if

$$A = [a_{ij}] \text{ and } X = \left[\begin{array}{c} \mathbf{v}_1 \\ \vdots \\ \mathbf{v}_n \end{array}\right], \text{ let } AX = \left[\begin{array}{c} \mathbf{u}_1 \\ \vdots \\ \mathbf{u}_n \end{array}\right]$$

where $\mathbf{u}_i = a_{i1}\mathbf{v}_1 + a_{i2}\mathbf{v}_2 + \cdots + a_{in}\mathbf{v}_n$ for each i.
Prove that:

a. $B(AX) = (BA)X$

b. $(A+A_1)X = AX + A_1X$

c. $A(X+X_1) = AX + AX_1$

d. $(kA)X = k(AX) = A(kX)$ if k is any number

e. $IX = X$ if I is the $n \times n$ identity matrix

f. Let E be an elementary matrix obtained by performing a row operation on the rows of I_n (see Section 2.5). Show that EX is the column resulting from performing that same row operation on the vectors (call them rows) of X. [*Hint*: Lemma 2.5.1.]

6.2 Subspaces and Spanning Sets

Chapter 5 is essentially about the subspaces of $\mathbb{R}^n$. We now extend this notion.

Definition 6.2 Subspaces of a Vector Space

If V is a vector space, a nonempty subset $U \subseteq V$ is called a ***subspace*** *of V if U is itself a vector space using the addition and scalar multiplication of V.*

Subspaces of $\mathbb{R}^n$ (as defined in Section 5.1) are subspaces in the present sense by Example 6.1.3. Moreover, the defining properties for a subspace of $\mathbb{R}^n$ actually *characterize* subspaces in general.

Theorem 6.2.1: Subspace Test

A subset U of a vector space is a subspace of V if and only if it satisfies the following three conditions:

1. *$\mathbf{0}$ lies in U where $\mathbf{0}$ is the zero vector of V.*
2. *If $\mathbf{u}_1$ and $\mathbf{u}_2$ are in U, then $\mathbf{u}_1 + \mathbf{u}_2$ is also in U.*
3. *If $\mathbf{u}$ is in U, then $a\mathbf{u}$ is also in U for each scalar a.*

Proof. If U is a subspace of V, then (2) and (3) hold by axioms A1 and S1 respectively, applied to the vector space U. Since U is nonempty (it is a vector space), choose $\mathbf{u}$ in U. Then (1) holds because $\mathbf{0} = 0\mathbf{u}$ is in U by (3) and Theorem 6.1.3.

Conversely, if (1), (2), and (3) hold, then axioms A1 and S1 hold because of (2) and (3), and axioms A2, A3, S2, S3, S4, and S5 hold in U because they hold in V. Axiom A4 holds because the zero vector $\mathbf{0}$ of V is actually in U by (1), and so serves as the zero of U. Finally, given $\mathbf{u}$ in U, then its negative $-\mathbf{u}$ in V is again in U by (3) because $-\mathbf{u} = (-1)\mathbf{u}$ (again using Theorem 6.1.3). Hence $-\mathbf{u}$ serves as the negative of $\mathbf{u}$ in U. □

Note that the proof of Theorem 6.2.1 shows that if U is a subspace of V, then U and V share the same zero vector, and that the negative of a vector in the space U is the same as its negative in V.

Example 6.2.1

If V is any vector space, show that $\{\mathbf{0}\}$ and V are subspaces of V.

Solution. $U = V$ clearly satisfies the conditions of the subspace test. As to $U = \{\mathbf{0}\}$, it satisfies the conditions because $\mathbf{0} + \mathbf{0} = \mathbf{0}$ and $a\mathbf{0} = \mathbf{0}$ for all a in $\mathbb{R}$.

The vector space $\{\mathbf{0}\}$ is called the **zero subspace** of V.

Example 6.2.2

Let $\mathbf{v}$ be a vector in a vector space V. Show that the set

$$\mathbb{R}\mathbf{v} = \{a\mathbf{v} \mid a \text{ in } \mathbb{R}\}$$

of all scalar multiples of $\mathbf{v}$ is a subspace of V.

Solution. Because $\mathbf{0} = 0\mathbf{v}$, it is clear that $\mathbf{0}$ lies in $\mathbb{R}\mathbf{v}$. Given two vectors $a\mathbf{v}$ and $a_1\mathbf{v}$ in $\mathbb{R}\mathbf{v}$, their sum $a\mathbf{v} + a_1\mathbf{v} = (a + a_1)\mathbf{v}$ is also a scalar multiple of $\mathbf{v}$ and so lies in $\mathbb{R}\mathbf{v}$. Hence $\mathbb{R}\mathbf{v}$ is closed under addition. Finally, given $a\mathbf{v}$, $r(a\mathbf{v}) = (ra)\mathbf{v}$ lies in $\mathbb{R}\mathbf{v}$ for all $r \in \mathbb{R}$, so $\mathbb{R}\mathbf{v}$ is closed under scalar multiplication. Hence the subspace test applies.

In particular, given $\mathbf{d} \neq \mathbf{0}$ in $\mathbb{R}^3$, $\mathbb{R}\mathbf{d}$ is the line through the origin with direction vector $\mathbf{d}$.

The space $\mathbb{R}\mathbf{v}$ in Example 6.2.2 is described by giving the *form* of each vector in $\mathbb{R}\mathbf{v}$. The next example describes a subset U of the space $\mathbf{M}_{nn}$ by giving a *condition* that each matrix of U must satisfy.

Example 6.2.3

Let A be a fixed matrix in $\mathbf{M}_{nn}$. Show that $U = \{X \text{ in } \mathbf{M}_{nn} \mid AX = XA\}$ is a subspace of $\mathbf{M}_{nn}$.

Solution. If 0 is the $n \times n$ zero matrix, then $A0 = 0A$, so 0 satisfies the condition for membership in U. Next suppose that X and X_1 lie in U so that $AX = XA$ and $AX_1 = X_1A$. Then

$$A(X + X_1) = AX + AX_1 = XA + X_1A + (X + X_1)A$$
$$A(aX) = a(AX) = a(XA) = (aX)A$$

for all a in $\mathbb{R}$, so both $X + X_1$ and aX lie in U. Hence U is a subspace of $\mathbf{M}_{nn}$.

Suppose $p(x)$ is a polynomial and a is a number. Then the number $p(a)$ obtained by replacing x by a in the expression for $p(x)$ is called the **evaluation** of $p(x)$ at a. For example, if $p(x) = 5 - 6x + 2x^2$, then the evaluation of $p(x)$ at $a = 2$ is $p(2) = 5 - 12 + 8 = 1$. If $p(a) = 0$, the number a is called a **root** of $p(x)$.

Example 6.2.4

Consider the set U of all polynomials in $\mathbf{P}$ that have 3 as a root:

$$U = \{p(x) \in \mathbf{P} \mid p(3) = 0\}$$

Show that U is a subspace of $\mathbf{P}$.

Solution. Clearly, the zero polynomial lies in U. Now let $p(x)$ and $q(x)$ lie in U so $p(3) = 0$ and $q(3) = 0$. We have $(p+q)(x) = p(x) + q(x)$ for all x, so $(p+q)(3) = p(3) + q(3) = 0 + 0 = 0$, and U is closed under addition. The verification that U is closed under scalar multiplication is similar.

Recall that the space $\mathbf{P}_n$ consists of all polynomials of the form

$$a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$$

where $a_0, a_1, a_2, \dots, a_n$ are real numbers, and so is closed under the addition and scalar multiplication in $\mathbf{P}$. Moreover, the zero polynomial is included in $\mathbf{P}_n$. Thus the subspace test gives Example 6.2.5.

Example 6.2.5

$\mathbf{P}_n$ is a subspace of $\mathbf{P}$ for each $n \geq 0$.

The next example involves the notion of the derivative f' of a function f. (If the reader is not familiar with calculus, this example may be omitted.) A function f defined on the interval $[a, b]$ is called **differentiable** if the derivative $f'(r)$ exists at every r in $[a, b]$.

Example 6.2.6

Show that the subset $\mathbf{D}[a, b]$ of all **differentiable functions** on $[a, b]$ is a subspace of the vector space $\mathbf{F}[a, b]$ of all functions on $[a, b]$.

Solution. The derivative of any constant function is the constant function 0; in particular, 0 itself is differentiable and so lies in $\mathbf{D}[a, b]$. If f and g both lie in $\mathbf{D}[a, b]$ (so that f' and g' exist), then it is a theorem of calculus that $f+g$ and rf are both differentiable for any $r \in \mathbb{R}$. In fact, $(f+g)' = f' + g'$ and $(rf)' = rf'$, so both lie in $\mathbf{D}[a, b]$. This shows that $\mathbf{D}[a, b]$ is a subspace of $\mathbf{F}[a, b]$.

Linear Combinations and Spanning Sets

Definition 6.3 Linear Combinations and Spanning

*Let $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ be a set of vectors in a vector space V. As in $\mathbb{R}^n$, a vector $\mathbf{v}$ is called a **linear combination** of the vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ if it can be expressed in the form*

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

*where $a_1, a_2, \dots, a_n$ are scalars, called the **coefficients** of $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$. The set of all linear combinations of these vectors is called their **span**, and is denoted by*

$$\text{span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\} = \{a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_n\mathbf{v}_n \mid a_i \text{ in } \mathbb{R}\}$$

If it happens that $V = \text{span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$, these vectors are called a **spanning set** for V. For example, the span of two vectors $\mathbf{v}$ and $\mathbf{w}$ is the set

$$\text{span}\{\mathbf{v}, \mathbf{w}\} = \{s\mathbf{v} + t\mathbf{w} \mid s \text{ and } t \text{ in } \mathbb{R}\}$$

of all sums of scalar multiples of these vectors.

Example 6.2.7

Consider the vectors $p_1 = 1 + x + 4x^2$ and $p_2 = 1 + 5x + x^2$ in $\mathbf{P}_2$. Determine whether p_1 and p_2 lie in $\text{span}\{1 + 2x - x^2, 3 + 5x + 2x^2\}$.

Solution. For p_1, we want to determine if s and t exist such that

$$p_1 = s(1 + 2x - x^2) + t(3 + 5x + 2x^2)$$

Equating coefficients of powers of x (where $x^0 = 1$) gives

$$1 = s + 3t, \quad 1 = 2s + 5t, \quad \text{and} \quad 4 = -s + 2t$$

These equations have the solution $s = -2$ and $t = 1$, so p_1 is indeed in $\text{span}\{1 + 2x - x^2, 3 + 5x + 2x^2\}$.
Turning to $p_2 = 1 + 5x + x^2$, we are looking for s and t such that

$$p_2 = s(1 + 2x - x^2) + t(3 + 5x + 2x^2)$$

Again equating coefficients of powers of x gives equations $1 = s + 3t$, $5 = 2s + 5t$, and $1 = -s + 2t$. But in this case there is no solution, so p_2 is *not* in $\text{span}\{1 + 2x - x^2, 3 + 5x + 2x^2\}$.

We saw in Example 5.1.6 that $\mathbb{R}^m = \text{span}\,\{\mathbf{e}_1,\ \mathbf{e}_2,\ \dots,\ \mathbf{e}_m\}$ where the vectors $\mathbf{e}_1,\ \mathbf{e}_2,\ \dots,\ \mathbf{e}_m$ are the columns of the $m \times m$ identity matrix. Of course $\mathbb{R}^m = \mathbf{M}_{m1}$ is the set of all $m \times 1$ matrices, and there is an analogous spanning set for each space $\mathbf{M}_{mn}$. For example, each 2×2 matrix has the form

$$\left[\begin{array}{cc} a & b \\ c & d \end{array}\right] = a\left[\begin{array}{cc} 1 & 0 \\ 0 & 0 \end{array}\right] + b\left[\begin{array}{cc} 0 & 1 \\ 0 & 0 \end{array}\right] + c\left[\begin{array}{cc} 0 & 0 \\ 1 & 0 \end{array}\right] + d\left[\begin{array}{cc} 0 & 0 \\ 0 & 1 \end{array}\right]$$

so

$$\mathbf{M}_{22} = \text{span}\left\{\left[\begin{array}{cc} 1 & 0 \\ 0 & 0 \end{array}\right],\ \left[\begin{array}{cc} 0 & 1 \\ 0 & 0 \end{array}\right],\ \left[\begin{array}{cc} 0 & 0 \\ 1 & 0 \end{array}\right],\ \left[\begin{array}{cc} 0 & 0 \\ 0 & 1 \end{array}\right]\right\}$$

Similarly, we obtain

Example 6.2.8

$\mathbf{M}_{mn}$ is the span of the set of all $m \times n$ matrices with exactly one entry equal to 1, and all other entries zero.

The fact that every polynomial in $\mathbf{P}_n$ has the form $a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$ where each a_i is in $\mathbb{R}$ shows that

Example 6.2.9

$\mathbf{P}_n = \text{span}\,\{1,\ x,\ x^2,\ \dots,\ x^n\}$.

In Example 6.2.2 we saw that $\text{span}\,\{\mathbf{v}\} = \{a\mathbf{v} \mid a \text{ in } \mathbb{R}\} = \mathbb{R}\mathbf{v}$ is a subspace for any vector $\mathbf{v}$ in a vector space V. More generally, the span of *any* set of vectors is a subspace. In fact, the proof of Theorem 5.1.1 goes through to prove:

Theorem 6.2.2

Let $U = \text{span}\,\{\mathbf{v}_1,\ \mathbf{v}_2,\ \dots,\ \mathbf{v}_n\}$ *in a vector space* V*. Then:*

1. U *is a subspace of* V *containing each of* $\mathbf{v}_1,\ \mathbf{v}_2,\ \dots,\ \mathbf{v}_n$.
2. U *is the "smallest" subspace containing these vectors in the sense that any subspace that contains each of* $\mathbf{v}_1,\ \mathbf{v}_2,\ \dots,\ \mathbf{v}_n$ *must contain* U.

Here is how condition 2 in Theorem 6.2.2 is used. Given vectors $\mathbf{v}_1,\ \dots,\ \mathbf{v}_k$ in a vector space V and a subspace $U \subseteq V$, then:

$$\text{span}\,\{\mathbf{v}_1,\ \dots,\ \mathbf{v}_n\} \subseteq U \Leftrightarrow \text{ each } \mathbf{v}_i \in U$$

The following examples illustrate this.

Example 6.2.10

Show that $\mathbf{P}_3 = \text{span}\,\{x^2 + x^3,\ x,\ 2x^2 + 1,\ 3\}$.

Solution. Write $U = \text{span}\,\{x^2 + x^3,\ x,\ 2x^2 + 1,\ 3\}$. Then $U \subseteq \mathbf{P}_3$, and we use the fact that

$\mathbf{P}_3 = \text{span}\{1, x, x^2, x^3\}$ to show that $\mathbf{P}_3 \subseteq U$. In fact, x and $1 = \frac{1}{3} \cdot 3$ clearly lie in U. But then successively,

$$x^2 = \tfrac{1}{2}[(2x^2+1)-1] \quad \text{and} \quad x^3 = (x^2+x^3)-x^2$$

also lie in U. Hence $\mathbf{P}_3 \subseteq U$ by Theorem 6.2.2.

Example 6.2.11

Let $\mathbf{u}$ and $\mathbf{v}$ be two vectors in a vector space V. Show that

$$\text{span}\{\mathbf{u}, \mathbf{v}\} = \text{span}\{\mathbf{u}+2\mathbf{v}, \mathbf{u}-\mathbf{v}\}$$

Solution. We have $\text{span}\{\mathbf{u}+2\mathbf{v}, \mathbf{u}-\mathbf{v}\} \subseteq \text{span}\{\mathbf{u}, \mathbf{v}\}$ by Theorem 6.2.2 because both $\mathbf{u}+2\mathbf{v}$ and $\mathbf{u}-\mathbf{v}$ lie in $\text{span}\{\mathbf{u}, \mathbf{v}\}$. On the other hand,

$$\mathbf{u} = \tfrac{1}{3}(\mathbf{u}+2\mathbf{v}) + \tfrac{2}{3}(\mathbf{u}-\mathbf{v}) \quad \text{and} \quad \mathbf{v} = \tfrac{1}{3}(\mathbf{u}+2\mathbf{v}) - \tfrac{1}{3}(\mathbf{u}-\mathbf{v})$$

so $\text{span}\{\mathbf{u}, \mathbf{v}\} \subseteq \text{span}\{\mathbf{u}+2\mathbf{v}, \mathbf{u}-\mathbf{v}\}$, again by Theorem 6.2.2.

Exercises for 6.2

Exercise 6.2.1 Which of the following are subspaces of $\mathbf{P}_3$? Support your answer.

a. $U = \{f(x) \mid f(x) \in \mathbf{P}_3,\ f(2) = 1\}$

b. $U = \{xg(x) \mid g(x) \in \mathbf{P}_2\}$

c. $U = \{xg(x) \mid g(x) \in \mathbf{P}_3\}$

d. $U = \{xg(x) + (1-x)h(x) \mid g(x) \text{ and } h(x) \in \mathbf{P}_2\}$

e. $U =$ The set of all polynomials in $\mathbf{P}_3$ with constant term 0

f. $U = \{f(x) \mid f(x) \in \mathbf{P}_3,\ \deg f(x) = 3\}$

Exercise 6.2.2 Which of the following are subspaces of $\mathbf{M}_{22}$? Support your answer.

a. $U = \left\{ \left[\begin{array}{cc} a & b \\ 0 & c \end{array} \right] \middle| \ a, b, \text{ and } c \text{ in } \mathbb{R} \right\}$

b. $U = \left\{ \left[\begin{array}{cc} a & b \\ c & d \end{array} \right] \middle| \ a+b=c+d;\ a, b, c, d \text{ in } \mathbb{R} \right\}$

c. $U = \{A \mid A \in \mathbf{M}_{22},\ A = A^T\}$

d. $U = \{A \mid A \in \mathbf{M}_{22},\ AB = 0\}$, B a fixed 2×2 matrix

e. $U = \{A \mid A \in \mathbf{M}_{22},\ A^2 = A\}$

f. $U = \{A \mid A \in \mathbf{M}_{22},\ A \text{ is not invertible}\}$

g. $U = \{A \mid A \in \mathbf{M}_{22},\ BAC = CAB\}$, B and C fixed 2×2 matrices

Exercise 6.2.3 Which of the following are subspaces of $\mathbf{F}[0, 1]$? Support your answer.

a. $U = \{f \mid f(0) = 0\}$

b. $U = \{f \mid f(0) = 1\}$

c. $U = \{f \mid f(0) = f(1)\}$

d. $U = \{f \mid f(x) \geq 0$ for all x in $[0,\ 1]\}$

e. $U = \{f \mid f(x) = f(y)$ for all x and y in $[0,\ 1]\}$

f. $U = \{f \mid f(x+y) = f(x) + f(y)$ for all x and y in $[0,\ 1]\}$

g. $U = \{f \mid f$ is integrable and $\int_0^1 f(x)dx = 0\}$

Exercise 6.2.4 Let A be an $m \times n$ matrix. For which columns $\mathbf{b}$ in $\mathbb{R}^m$ is $U = \{\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n,\ A\mathbf{x} = \mathbf{b}\}$ a subspace of $\mathbb{R}^n$? Support your answer.

Exercise 6.2.5 Let $\mathbf{x}$ be a vector in $\mathbb{R}^n$ (written as a column), and define $U = \{A\mathbf{x} \mid A \in \mathbf{M}_{mn}\}$.

a. Show that U is a subspace of $\mathbb{R}^m$.

b. Show that $U = \mathbb{R}^m$ if $\mathbf{x} \neq \mathbf{0}$.

Exercise 6.2.6 Write each of the following as a linear combination of $x+1$, x^2+x, and x^2+2.

a. x^2+3x+2

b. $2x^2-3x+1$

c. x^2+1

d. x

Exercise 6.2.7 Determine whether $\mathbf{v}$ lies in span $\{\mathbf{u},\ \mathbf{w}\}$ in each case.

a. $\mathbf{v} = 3x^2-2x-1$; $\mathbf{u} = x^2+1$, $\mathbf{w} = x+2$

b. $\mathbf{v} = x$; $\mathbf{u} = x^2+1$, $\mathbf{w} = x+2$

c. $\mathbf{v} = \left[\begin{array}{rr} 1 & 3 \\ -1 & 1 \end{array}\right]$; $\mathbf{u} = \left[\begin{array}{rr} 1 & -1 \\ 2 & 1 \end{array}\right]$, $\mathbf{w} = \left[\begin{array}{rr} 2 & 1 \\ 1 & 0 \end{array}\right]$

d. $\mathbf{v} = \left[\begin{array}{rr} 1 & -4 \\ 5 & 3 \end{array}\right]$; $\mathbf{u} = \left[\begin{array}{rr} 1 & -1 \\ 2 & 1 \end{array}\right]$, $\mathbf{w} = \left[\begin{array}{rr} 2 & 1 \\ 1 & 0 \end{array}\right]$

Exercise 6.2.8 Which of the following functions lie in span $\{\cos^2 x,\ \sin^2 x\}$? (Work in $\mathbf{F}[0,\ \pi]$.)

a. $\cos 2x$

b. 1

c. x^2

d. $1+x^2$

Exercise 6.2.9

a. Show that $\mathbb{R}^3$ is spanned by $\{(1,\ 0,\ 1),\ (1,\ 1,\ 0),\ (0,\ 1,\ 1)\}$.

b. Show that $\mathbf{P}_2$ is spanned by $\{1+2x^2,\ 3x,\ 1+x\}$.

c. Show that $\mathbf{M}_{22}$ is spanned by $\left\{\left[\begin{array}{cc} 1 & 0 \\ 0 & 0 \end{array}\right], \left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array}\right], \left[\begin{array}{cc} 0 & 1 \\ 1 & 0 \end{array}\right], \left[\begin{array}{cc} 1 & 1 \\ 0 & 1 \end{array}\right]\right\}$.

Exercise 6.2.10 If X and Y are two sets of vectors in a vector space V, and if $X \subseteq Y$, show that span $X \subseteq$ span Y.

Exercise 6.2.11 Let $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ denote vectors in a vector space V. Show that:

a. span $\{\mathbf{u},\ \mathbf{v},\ \mathbf{w}\}$ = span $\{\mathbf{u}+\mathbf{v},\ \mathbf{u}+\mathbf{w},\ \mathbf{v}+\mathbf{w}\}$

b. span $\{\mathbf{u},\ \mathbf{v},\ \mathbf{w}\}$ = span $\{\mathbf{u}-\mathbf{v},\ \mathbf{u}+\mathbf{w},\ \mathbf{w}\}$

Exercise 6.2.12 Show that

$$\text{span}\,\{\mathbf{v}_1,\ \mathbf{v}_2,\ \dots,\ \mathbf{v}_n,\ \mathbf{0}\} = \text{span}\,\{\mathbf{v}_1,\ \mathbf{v}_2,\ \dots,\ \mathbf{v}_n\}$$

holds for any set of vectors $\{\mathbf{v}_1,\ \mathbf{v}_2,\ \dots,\ \mathbf{v}_n\}$.

Exercise 6.2.13 If X and Y are nonempty subsets of a vector space V such that span X = span $Y = V$, must there be a vector common to both X and Y? Justify your answer.

Exercise 6.2.14 Is it possible that $\{(1,\ 2,\ 0),\ (1,\ 1,\ 1)\}$ can span the subspace $U = \{(a,\ b,\ 0) \mid a$ and b in $\mathbb{R}\}$?

Exercise 6.2.15 Describe span $\{\mathbf{0}\}$.

Exercise 6.2.16 Let $\mathbf{v}$ denote any vector in a vector space V. Show that span $\{\mathbf{v}\}$ = span $\{a\mathbf{v}\}$ for any $a \neq 0$.

Exercise 6.2.17 Determine all subspaces of $\mathbb{R}\mathbf{v}$ where $\mathbf{v} \neq \mathbf{0}$ in some vector space V.

Exercise 6.2.18 Suppose V = span $\{\mathbf{v}_1,\ \mathbf{v}_2,\ \dots,\ \mathbf{v}_n\}$. If $\mathbf{u} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_n\mathbf{v}_n$ where the a_i are in $\mathbb{R}$ and $a_1 \neq 0$, show that V = span $\{\mathbf{u},\ \mathbf{v}_2,\ \dots,\ \mathbf{v}_n\}$.

Exercise 6.2.19 If $\mathbf{M}_{nn}$ = span $\{A_1,\ A_2,\ \dots,\ A_k\}$, show that $\mathbf{M}_{nn}$ = span $\{A_1^T,\ A_2^T,\ \dots,\ A_k^T\}$.

Exercise 6.2.20 If $\mathbf{P}_n$ = span $\{p_1(x),\ p_2(x),\ \dots,\ p_k(x)\}$ and a is in $\mathbb{R}$, show that $p_i(a) \neq 0$ for some i.

Exercise 6.2.21 Let U be a subspace of a vector space V.

a. If $a\mathbf{u}$ is in U where $a \neq 0$, show that $\mathbf{u}$ is in U.

b. If $\mathbf{u}$ and $\mathbf{u}+\mathbf{v}$ are in U, show that $\mathbf{v}$ is in U.

Exercise 6.2.22 Let U be a nonempty subset of a vector space V. Show that U is a subspace of V if and only if $\mathbf{u}_1 + a\mathbf{u}_2$ lies in U for all $\mathbf{u}_1$ and $\mathbf{u}_2$ in U and all a in $\mathbb{R}$.

Exercise 6.2.23 Let $U = \{p(x) \text{ in } \mathbf{P} \mid p(3) = 0\}$ be the set in Example 6.2.4. Use the factor theorem (see Section 6.5) to show that U consists of multiples of $x-3$; that is, show that $U = \{(x-3)q(x) \mid q(x) \in \mathbf{P}\}$. Use this to show that U is a subspace of $\mathbf{P}$.

Exercise 6.2.24 Let $A_1, A_2, \dots, A_m$ denote $n \times n$ matrices. If $\mathbf{0} \neq \mathbf{y} \in \mathbb{R}^n$ and $A_1\mathbf{y} = A_2\mathbf{y} = \cdots = A_m\mathbf{y} = \mathbf{0}$, show that $\{A_1, A_2, \dots, A_m\}$ cannot span $\mathbf{M}_{nn}$.

Exercise 6.2.25 Let $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ and $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n\}$ be sets of vectors in a vector space, and let

$$X = \left[\begin{array}{c} \mathbf{v}_1 \\ \vdots \\ \mathbf{v}_n \end{array}\right] \quad Y = \left[\begin{array}{c} \mathbf{u}_1 \\ \vdots \\ \mathbf{u}_n \end{array}\right]$$

as in Exercise 6.1.18.

a. Show that $\text{span}\,\{\mathbf{v}_1, \dots, \mathbf{v}_n\} \subseteq \text{span}\,\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ if and only if $AY = X$ for some $n \times n$ matrix A.

b. If $X = AY$ where A is invertible, show that $\text{span}\,\{\mathbf{v}_1, \dots, \mathbf{v}_n\} = \text{span}\,\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$.

Exercise 6.2.26 If U and W are subspaces of a vector space V, let $U \cup W = \{\mathbf{v} \mid \mathbf{v} \text{ is in } U \text{ or } \mathbf{v} \text{ is in } W\}$. Show that $U \cup W$ is a subspace if and only if $U \subseteq W$ or $W \subseteq U$.

Exercise 6.2.27 Show that $\mathbf{P}$ cannot be spanned by a finite set of polynomials.

6.3 Linear Independence and Dimension

Definition 6.4 Linear Independence and Dependence

As in $\mathbb{R}^n$, a set of vectors $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ in a vector space V is called ***linearly independent*** *(or simply* ***independent****) if it satisfies the following condition:*

$$\text{If} \quad s_1\mathbf{v}_1 + s_2\mathbf{v}_2 + \cdots + s_n\mathbf{v}_n = \mathbf{0}, \quad \text{then} \quad s_1 = s_2 = \cdots = s_n = 0.$$

A set of vectors that is not linearly independent is said to be ***linearly dependent*** *(or simply* ***dependent****).*

The **trivial linear combination** of the vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ is the one with every coefficient zero:

$$0\mathbf{v}_1 + 0\mathbf{v}_2 + \cdots + 0\mathbf{v}_n$$

This is obviously one way of expressing $\mathbf{0}$ as a linear combination of the vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$, and they are linearly independent when it is the *only* way.

Example 6.3.1

Show that $\{1+x,\ 3x+x^2,\ 2+x-x^2\}$ is independent in $\mathbf{P}_2$.

Solution. Suppose a linear combination of these polynomials vanishes.

$$s_1(1+x) + s_2(3x+x^2) + s_3(2+x-x^2) = 0$$

Equating the coefficients of 1, x, and x^2 gives a set of linear equations.

$$\begin{aligned} s_1 + + 2s_3 &= 0 \\ s_1 + 3s_2 + s_3 &= 0 \\ s_2 - s_3 &= 0 \end{aligned}$$

The only solution is $s_1 = s_2 = s_3 = 0$.

Example 6.3.2

Show that $\{\sin x,\ \cos x\}$ is independent in the vector space $\mathbf{F}[0,\ 2\pi]$ of functions defined on the interval $[0,\ 2\pi]$.

Solution. Suppose that a linear combination of these functions vanishes.

$$s_1(\sin x) + s_2(\cos x) = 0$$

This must hold for *all* values of x in $[0,\ 2\pi]$ (by the definition of equality in $\mathbf{F}[0,\ 2\pi]$). Taking $x = 0$ yields $s_2 = 0$ (because $\sin 0 = 0$ and $\cos 0 = 1$). Similarly, $s_1 = 0$ follows from taking $x = \frac{\pi}{2}$ (because $\sin\frac{\pi}{2} = 1$ and $\cos\frac{\pi}{2} = 0$).

Example 6.3.3

Suppose that $\{\mathbf{u},\ \mathbf{v}\}$ is an independent set in a vector space V. Show that $\{\mathbf{u} + 2\mathbf{v},\ \mathbf{u} - 3\mathbf{v}\}$ is also independent.

Solution. Suppose a linear combination of $\mathbf{u} + 2\mathbf{v}$ and $\mathbf{u} - 3\mathbf{v}$ vanishes:

$$s(\mathbf{u} + 2\mathbf{v}) + t(\mathbf{u} - 3\mathbf{v}) = \mathbf{0}$$

We must deduce that $s = t = 0$. Collecting terms involving $\mathbf{u}$ and $\mathbf{v}$ gives

$$(s+t)\mathbf{u} + (2s - 3t)\mathbf{v} = \mathbf{0}$$

Because $\{\mathbf{u},\ \mathbf{v}\}$ is independent, this yields linear equations $s + t = 0$ and $2s - 3t = 0$. The only solution is $s = t = 0$.

Example 6.3.4

Show that any set of polynomials of distinct degrees is independent.

Solution. Let $p_1,\ p_2,\ \dots,\ p_m$ be polynomials where $\deg(p_i) = d_i$. By relabelling if necessary, we may assume that $d_1 > d_2 > \cdots > d_m$. Suppose that a linear combination vanishes:

$$t_1p_1 + t_2p_2 + \cdots + t_mp_m = 0$$

where each t_i is in $\mathbb{R}$. As $\deg(p_1) = d_1$, let ax^{d_1} be the term in p_1 of highest degree, where $a \neq 0$.

Since $d_1 > d_2 > \cdots > d_m$, it follows that $t_1 a x^{d_1}$ is the only term of degree d_1 in the linear combination $t_1 p_1 + t_2 p_2 + \cdots + t_m p_m = 0$. This means that $t_1 a x^{d_1} = 0$, whence $t_1 a = 0$, hence $t_1 = 0$ (because $a \neq 0$). But then $t_2 p_2 + \cdots + t_m p_m = 0$ so we can repeat the argument to show that $t_2 = 0$. Continuing, we obtain $t_i = 0$ for each i, as desired.

Example 6.3.5

Suppose that A is an $n \times n$ matrix such that $A^k = 0$ but $A^{k-1} \neq 0$. Show that $B = \{I,\ A,\ A^2,\ \ldots,\ A^{k-1}\}$ is independent in $\mathbf{M}_{nn}$.

Solution. Suppose $r_0 I + r_1 A + r_2 A^2 + \cdots + r_{k-1} A^{k-1} = 0$. Multiply by A^{k-1}:

$$r_0 A^{k-1} + r_1 A^k + r_2 A^{k+1} + \cdots + r_{k-1} A^{2k-2} = 0$$

Since $A^k = 0$, all the higher powers are zero, so this becomes $r_0 A^{k-1} = 0$. But $A^{k-1} \neq 0$, so $r_0 = 0$, and we have $r_1 A^1 + r_2 A^2 + \cdots + r_{k-1} A^{k-1} = 0$. Now multiply by A^{k-2} to conclude that $r_1 = 0$. Continuing, we obtain $r_i = 0$ for each i, so B is independent.

The next example collects several useful properties of independence for reference.

Example 6.3.6

Let V denote a vector space.

1. If $\mathbf{v} \neq \mathbf{0}$ in V, then $\{\mathbf{v}\}$ is an independent set.
2. No independent set of vectors in V can contain the zero vector.

Solution.

1. Let $t\mathbf{v} = \mathbf{0}$, t in $\mathbb{R}$. If $t \neq 0$, then $\mathbf{v} = 1\mathbf{v} = \frac{1}{t}(t\mathbf{v}) = \frac{1}{t}\mathbf{0} = \mathbf{0}$, contrary to assumption. So $t = 0$.
2. If $\{\mathbf{v}_1,\ \mathbf{v}_2,\ \ldots,\ \mathbf{v}_k\}$ is independent and (say) $\mathbf{v}_2 = \mathbf{0}$, then $0\mathbf{v}_1 + 1\mathbf{v}_2 + \cdots + 0\mathbf{v}_k = \mathbf{0}$ is a nontrivial linear combination that vanishes, contrary to the independence of $\{\mathbf{v}_1,\ \mathbf{v}_2,\ \ldots,\ \mathbf{v}_k\}$.

A set of vectors is independent if $\mathbf{0}$ is a linear combination in a unique way. The following theorem shows that *every* linear combination of these vectors has uniquely determined coefficients, and so extends Theorem 5.2.1.

Theorem 6.3.1

Let $\{\mathbf{v}_1,\ \mathbf{v}_2,\ \ldots,\ \mathbf{v}_n\}$ be a linearly independent set of vectors in a vector space V. If a vector $\mathbf{v}$ has *two (ostensibly different) representations*

$$\mathbf{v} = s_1\mathbf{v}_1 + s_2\mathbf{v}_2 + \cdots + s_n\mathbf{v}_n$$
$$\mathbf{v} = t_1\mathbf{v}_1 + t_2\mathbf{v}_2 + \cdots + t_n\mathbf{v}_n$$

as linear combinations of these vectors, then $s_1 = t_1,\ s_2 = t_2,\ \dots,\ s_n = t_n$*. In other words, every vector in* V *can be written in a unique way as a linear combination of the* $\mathbf{v}_i$*.*

Proof. Subtracting the equations given in the theorem gives

$$(s_1 - t_1)\mathbf{v}_1 + (s_2 - t_2)\mathbf{v}_2 + \cdots + (s_n - t_n)\mathbf{v}_n = \mathbf{0}$$

The independence of $\{\mathbf{v}_1,\ \mathbf{v}_2,\ \dots,\ \mathbf{v}_n\}$ gives $s_i - t_i = 0$ for each i, as required. □

The following theorem extends (and proves) Theorem 5.2.4, and is one of the most useful results in linear algebra.

Theorem 6.3.2: Fundamental Theorem

Suppose a vector space V *can be spanned by* n *vectors. If any set of* m *vectors in* V *is linearly independent, then* $m \leq n$*.*

Proof. Let $V = \text{span}\{\mathbf{v}_1,\ \mathbf{v}_2,\ \dots,\ \mathbf{v}_n\}$, and suppose that $\{\mathbf{u}_1,\ \mathbf{u}_2,\ \dots,\ \mathbf{u}_m\}$ is an independent set in V. Then $\mathbf{u}_1 = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_n\mathbf{v}_n$ where each a_i is in $\mathbb{R}$. As $\mathbf{u}_1 \neq \mathbf{0}$ (Example 6.3.6), not all of the a_i are zero, say $a_1 \neq 0$ (after relabelling the $\mathbf{v}_i$). Then $V = \text{span}\{\mathbf{u}_1,\ \mathbf{v}_2,\ \mathbf{v}_3,\ \dots,\ \mathbf{v}_n\}$ as the reader can verify. Hence, write $\mathbf{u}_2 = b_1\mathbf{u}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 + \cdots + c_n\mathbf{v}_n$. Then some $c_i \neq 0$ because $\{\mathbf{u}_1,\ \mathbf{u}_2\}$ is independent; so, as before, $V = \text{span}\{\mathbf{u}_1,\ \mathbf{u}_2,\ \mathbf{v}_3,\ \dots,\ \mathbf{v}_n\}$, again after possible relabelling of the $\mathbf{v}_i$. If $m > n$, this procedure continues until all the vectors $\mathbf{v}_i$ are replaced by the vectors $\mathbf{u}_1,\ \mathbf{u}_2,\ \dots,\ \mathbf{u}_n$. In particular, $V = \text{span}\{\mathbf{u}_1,\ \mathbf{u}_2,\ \dots,\ \mathbf{u}_n\}$. But then $\mathbf{u}_{n+1}$ is a linear combination of $\mathbf{u}_1,\ \mathbf{u}_2,\ \dots,\ \mathbf{u}_n$ contrary to the independence of the $\mathbf{u}_i$. Hence, the assumption $m > n$ cannot be valid, so $m \leq n$ and the theorem is proved. □

If $V = \text{span}\{\mathbf{v}_1,\ \mathbf{v}_2,\ \dots,\ \mathbf{v}_n\}$, and if $\{\mathbf{u}_1,\ \mathbf{u}_2,\ \dots,\ \mathbf{u}_m\}$ is an independent set in V, the above proof shows not only that $m \leq n$ but also that m of the (spanning) vectors $\mathbf{v}_1,\ \mathbf{v}_2,\ \dots,\ \mathbf{v}_n$ can be replaced by the (independent) vectors $\mathbf{u}_1,\ \mathbf{u}_2,\ \dots,\ \mathbf{u}_m$ and the resulting set will still span V. In this form the result is called the **Steinitz Exchange Lemma**.

Definition 6.5 Basis of a Vector Space

*As in $\mathbb{R}^n$, a set $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ of vectors in a vector space V is called a **basis** of V if it satisfies the following two conditions:*

1. $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ *is linearly independent*
2. $V = \text{span}\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$

Thus if a set of vectors $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ is a basis, then *every* vector in V can be written as a linear combination of these vectors in a *unique* way (Theorem 6.3.1). But even more is true: Any two (finite) bases of V contain the same number of vectors.

Theorem 6.3.3: Invariance Theorem

Let $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ and $\{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_m\}$ be two bases of a vector space V. Then $n = m$.

Proof. Because $V = \text{span}\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ and $\{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_m\}$ is independent, it follows from Theorem 6.3.2 that $m \leq n$. Similarly $n \leq m$, so $n = m$, as asserted. □

Theorem 6.3.3 guarantees that no matter which basis of V is chosen it contains the same number of vectors as any other basis. Hence there is no ambiguity about the following definition.

Definition 6.6 Dimension of a Vector Space

*If $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ is a basis of the nonzero vector space V, the number n of vectors in the basis is called the **dimension** of V, and we write*

$$\dim V = n$$

The zero vector space $\{\mathbf{0}\}$ is defined to have dimension 0:

$$\dim\{\mathbf{0}\} = 0$$

In our discussion to this point we have always assumed that a basis is nonempty and hence that the dimension of the space is at least 1. However, the zero space $\{\mathbf{0}\}$ has *no* basis (by Example 6.3.6) so our insistence that $\dim\{\mathbf{0}\} = 0$ amounts to saying that the *empty* set of vectors is a basis of $\{\mathbf{0}\}$. Thus the statement that "the dimension of a vector space is the number of vectors in any basis" holds even for the zero space.

We saw in Example 5.2.9 that $\dim(\mathbb{R}^n) = n$ and, if $\mathbf{e}_j$ denotes column j of I_n, that $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ is a basis (called the standard basis). In Example 6.3.7 below, similar considerations apply to the space $\mathbf{M}_{mn}$ of all $m \times n$ matrices; the verifications are left to the reader.

Example 6.3.7

The space $\mathbf{M}_{mn}$ has dimension mn, and one basis consists of all $m \times n$ matrices with exactly one entry equal to 1 and all other entries equal to 0. We call this the **standard basis** of $\mathbf{M}_{mn}$.

Example 6.3.8

Show that $\dim \mathbf{P}_n = n+1$ and that $\{1, x, x^2, \dots, x^n\}$ is a basis, called the **standard basis** of $\mathbf{P}_n$.

Solution. Each polynomial $p(x) = a_0 + a_1x + \cdots + a_nx^n$ in $\mathbf{P}_n$ is clearly a linear combination of $1, x, \dots, x^n$, so $\mathbf{P}_n = \text{span}\{1, x, \dots, x^n\}$. However, if a linear combination of these vectors vanishes, $a_0 1 + a_1 x + \cdots + a_n x^n = 0$, then $a_0 = a_1 = \cdots = a_n = 0$ because x is an indeterminate. So $\{1, x, \dots, x^n\}$ is linearly independent and hence is a basis containing $n+1$ vectors. Thus, $\dim(\mathbf{P}_n) = n+1$.

Example 6.3.9

If $\mathbf{v} \neq \mathbf{0}$ is any nonzero vector in a vector space V, show that $\text{span}\{\mathbf{v}\} = \mathbb{R}\mathbf{v}$ has dimension 1.

Solution. $\{\mathbf{v}\}$ clearly spans $\mathbb{R}\mathbf{v}$, and it is linearly independent by Example 6.3.6. Hence $\{\mathbf{v}\}$ is a basis of $\mathbb{R}\mathbf{v}$, and so $\dim \mathbb{R}\mathbf{v} = 1$.

Example 6.3.10

Let $A = \left[\begin{array}{cc} 1 & 1 \\ 0 & 0 \end{array} \right]$ and consider the subspace

$$U = \{X \text{ in } \mathbf{M}_{22} \mid AX = XA\}$$

of $\mathbf{M}_{22}$. Show that $\dim U = 2$ and find a basis of U.

Solution. It was shown in Example 6.2.3 that U is a subspace for any choice of the matrix A. In the present case, if $X = \left[\begin{array}{cc} x & y \\ z & w \end{array} \right]$ is in U, the condition $AX = XA$ gives $z = 0$ and $x = y + w$. Hence each matrix X in U can be written

$$X = \left[\begin{array}{cc} y+w & y \\ 0 & w \end{array} \right] = y \left[\begin{array}{cc} 1 & 1 \\ 0 & 0 \end{array} \right] + w \left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array} \right]$$

so $U = \text{span}\, B$ where $B = \left\{ \left[\begin{array}{cc} 1 & 1 \\ 0 & 0 \end{array} \right], \left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array} \right] \right\}$. Moreover, the set B is linearly independent (verify this), so it is a basis of U and $\dim U = 2$.

Example 6.3.11

Show that the set V of all symmetric 2×2 matrices is a vector space, and find the dimension of V.

Solution. A matrix A is symmetric if $A^T = A$. If A and B lie in V, then

$$(A+B)^T = A^T + B^T = A + B \quad \text{and} \quad (kA)^T = kA^T = kA$$

using Theorem 2.1.2. Hence $A+B$ and kA are also symmetric. As the 2×2 zero matrix is also in

V, this shows that V is a vector space (being a subspace of $\mathbf{M}_{22}$). Now a matrix A is symmetric when entries directly across the main diagonal are equal, so each 2×2 symmetric matrix has the form

$$\begin{bmatrix} a & c \\ c & b \end{bmatrix} = a\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + b\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} + c\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

Hence the set $B = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \right\}$ spans V, and the reader can verify that B is linearly independent. Thus B is a basis of V, so $\dim V = 3$.

It is frequently convenient to alter a basis by multiplying each basis vector by a nonzero scalar. The next example shows that this always produces another basis. The proof is left as Exercise 6.3.22.

Example 6.3.12

Let $B = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ be nonzero vectors in a vector space V. Given nonzero scalars $a_1, a_2, \dots, a_n$, write $D = \{a_1\mathbf{v}_1, a_2\mathbf{v}_2, \dots, a_n\mathbf{v}_n\}$. If B is independent or spans V, the same is true of D. In particular, if B is a basis of V, so also is D.

Exercises for 6.3

Exercise 6.3.1 Show that each of the following sets of vectors is independent.

a. $\{1+x, 1-x, x+x^2\}$ in $\mathbf{P}_2$

b. $\{x^2, x+1, 1-x-x^2\}$ in $\mathbf{P}_2$

c. $\left\{ \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & -1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} \right\}$ in $\mathbf{M}_{22}$

d. $\left\{ \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \right\}$ in $\mathbf{M}_{22}$

Exercise 6.3.2 Which of the following subsets of V are independent?

a. $V = \mathbf{P}_2$; $\{x^2+1, x+1, x\}$

b. $V = \mathbf{P}_2$; $\{x^2-x+3, 2x^2+x+5, x^2+5x+1\}$

c. $V = \mathbf{M}_{22}$; $\left\{ \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right\}$

d. $V = \mathbf{M}_{22}$; $\left\{ \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix} \right\}$

e. $V = \mathbf{F}[1, 2]$; $\left\{ \frac{1}{x}, \frac{1}{x^2}, \frac{1}{x^3} \right\}$

f. $V = \mathbf{F}[0, 1]$; $\left\{ \frac{1}{x^2+x-6}, \frac{1}{x^2-5x+6}, \frac{1}{x^2-9} \right\}$

Exercise 6.3.3 Which of the following are independent in $\mathbf{F}[0, 2\pi]$?

a. $\{\sin^2 x, \cos^2 x\}$

b. $\{1, \sin^2 x, \cos^2 x\}$

c. $\{x, \sin^2 x, \cos^2 x\}$

Exercise 6.3.4 Find all values of a such that the following are independent in $\mathbb{R}^3$.

a. $\{(1, -1, 0), (a, 1, 0), (0, 2, 3)\}$

b. $\{(2, a, 1), (1, 0, 1), (0, 1, 3)\}$

Exercise 6.3.5 Show that the following are bases of the space V indicated.

a. $\{(1, 1, 0), (1, 0, 1), (0, 1, 1)\}$; $V = \mathbb{R}^3$

b. $\{(-1,\ 1,\ 1),\ (1,\ -1,\ 1),\ (1,\ 1,\ -1)\}$; $V = \mathbb{R}^3$

c. $\left\{ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \right\}$; $V = \mathbf{M}_{22}$

d. $\{1+x,\ x+x^2,\ x^2+x^3,\ x^3\}$; $V = \mathbf{P}_3$

Exercise 6.3.6 Exhibit a basis and calculate the dimension of each of the following subspaces of $\mathbf{P}_2$.

a. $\{a(1+x)+b(x+x^2) \mid a \text{ and } b \text{ in } \mathbb{R}\}$

b. $\{a+b(x+x^2) \mid a \text{ and } b \text{ in } \mathbb{R}\}$

c. $\{p(x) \mid p(1)=0\}$

d. $\{p(x) \mid p(x)=p(-x)\}$

Exercise 6.3.7 Exhibit a basis and calculate the dimension of each of the following subspaces of $\mathbf{M}_{22}$.

a. $\{A \mid A^T = -A\}$

b. $\left\{ A \,\middle|\, A \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix} A \right\}$

c. $\left\{ A \,\middle|\, A \begin{bmatrix} 1 & 0 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \right\}$

d. $\left\{ A \,\middle|\, A \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 1 \end{bmatrix} A \right\}$

Exercise 6.3.8 Let $A = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$ and define $U = \{X \mid X \in \mathbf{M}_{22} \text{ and } AX = X\}$.

a. Find a basis of U containing A.

b. Find a basis of U not containing A.

Exercise 6.3.9 Show that the set $\mathbb{C}$ of all complex numbers is a vector space with the usual operations, and find its dimension.

Exercise 6.3.10

a. Let V denote the set of all 2×2 matrices with equal column sums. Show that V is a subspace of $\mathbf{M}_{22}$, and compute $\dim V$.

b. Repeat part (a) for 3×3 matrices.

c. Repeat part (a) for $n \times n$ matrices.

Exercise 6.3.11

a. Let $V = \{(x^2+x+1)p(x) \mid p(x) \text{ in } \mathbf{P}_2\}$. Show that V is a subspace of $\mathbf{P}_4$ and find $\dim V$. [*Hint*: If $f(x)g(x) = 0$ in $\mathbf{P}$, then $f(x) = 0$ or $g(x) = 0$.]

b. Repeat with $V = \{(x^2-x)p(x) \mid p(x) \text{ in } \mathbf{P}_3\}$, a subset of $\mathbf{P}_5$.

c. Generalize.

Exercise 6.3.12 In each case, either prove the assertion or give an example showing that it is false.

a. Every set of four nonzero polynomials in $\mathbf{P}_3$ is a basis.

b. $\mathbf{P}_2$ has a basis of polynomials $f(x)$ such that $f(0) = 0$.

c. $\mathbf{P}_2$ has a basis of polynomials $f(x)$ such that $f(0) = 1$.

d. Every basis of $\mathbf{M}_{22}$ contains a noninvertible matrix.

e. No independent subset of $\mathbf{M}_{22}$ contains a matrix A with $A^2 = 0$.

f. If $\{\mathbf{u},\ \mathbf{v},\ \mathbf{w}\}$ is independent then, $a\mathbf{u}+b\mathbf{v}+c\mathbf{w} = \mathbf{0}$ for some a, b, c.

g. $\{\mathbf{u},\ \mathbf{v},\ \mathbf{w}\}$ is independent if $a\mathbf{u}+b\mathbf{v}+c\mathbf{w} = \mathbf{0}$ for some a, b, c.

h. If $\{\mathbf{u},\ \mathbf{v}\}$ is independent, so is $\{\mathbf{u},\ \mathbf{u}+\mathbf{v}\}$.

i. If $\{\mathbf{u},\ \mathbf{v}\}$ is independent, so is $\{\mathbf{u},\ \mathbf{v},\ \mathbf{u}+\mathbf{v}\}$.

j. If $\{\mathbf{u},\ \mathbf{v},\ \mathbf{w}\}$ is independent, so is $\{\mathbf{u},\ \mathbf{v}\}$.

k. If $\{\mathbf{u},\ \mathbf{v},\ \mathbf{w}\}$ is independent, so is $\{\mathbf{u}+\mathbf{w},\ \mathbf{v}+\mathbf{w}\}$.

l. If $\{\mathbf{u},\ \mathbf{v},\ \mathbf{w}\}$ is independent, so is $\{\mathbf{u}+\mathbf{v}+\mathbf{w}\}$.

m. If $\mathbf{u} \neq \mathbf{0}$ and $\mathbf{v} \neq \mathbf{0}$ then $\{\mathbf{u}, \mathbf{v}\}$ is dependent if and only if one is a scalar multiple of the other.

n. If $\dim V = n$, then no set of more than n vectors can be independent.

o. If $\dim V = n$, then no set of fewer than n vectors can span V.

Exercise 6.3.13 Let $A \neq 0$ and $B \neq 0$ be $n \times n$ matrices, and assume that A is symmetric and B is skew-symmetric (that is, $B^T = -B$). Show that $\{A, B\}$ is independent.

Exercise 6.3.14 Show that every set of vectors containing a dependent set is again dependent.

Exercise 6.3.15 Show that every nonempty subset of an independent set of vectors is again independent.

Exercise 6.3.16 Let f and g be functions on $[a, b]$, and assume that $f(a) = 1 = g(b)$ and $f(b) = 0 = g(a)$. Show that $\{f, g\}$ is independent in $\mathbf{F}[a, b]$.

Exercise 6.3.17 Let $\{A_1, A_2, \dots, A_k\}$ be independent in $\mathbf{M}_{mn}$, and suppose that U and V are invertible matrices of size $m \times m$ and $n \times n$, respectively. Show that $\{UA_1V, UA_2V, \dots, UA_kV\}$ is independent.

Exercise 6.3.18 Show that $\{\mathbf{v}, \mathbf{w}\}$ is independent if and only if neither $\mathbf{v}$ nor $\mathbf{w}$ is a scalar multiple of the other.

Exercise 6.3.19 Assume that $\{\mathbf{u}, \mathbf{v}\}$ is independent in a vector space V. Write $\mathbf{u}' = a\mathbf{u} + b\mathbf{v}$ and $\mathbf{v}' = c\mathbf{u} + d\mathbf{v}$, where a, b, c, and d are numbers. Show that $\{\mathbf{u}', \mathbf{v}'\}$ is independent if and only if the matrix $\left[\begin{array}{cc} a & c \\ b & d \end{array}\right]$ is invertible. [*Hint*: Theorem 2.4.5.]

Exercise 6.3.20 If $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ is independent and $\mathbf{w}$ is not in $\text{span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$, show that:

a. $\{\mathbf{w}, \mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ is independent.

b. $\{\mathbf{v}_1 + \mathbf{w}, \mathbf{v}_2 + \mathbf{w}, \dots, \mathbf{v}_k + \mathbf{w}\}$ is independent.

Exercise 6.3.21 If $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ is independent, show that $\{\mathbf{v}_1, \mathbf{v}_1 + \mathbf{v}_2, \dots, \mathbf{v}_1 + \mathbf{v}_2 + \cdots + \mathbf{v}_k\}$ is also independent.

Exercise 6.3.22 Prove Example 6.3.12.

Exercise 6.3.23 Let $\{\mathbf{u}, \mathbf{v}, \mathbf{w}, \mathbf{z}\}$ be independent. Which of the following are dependent?

a. $\{\mathbf{u} - \mathbf{v}, \mathbf{v} - \mathbf{w}, \mathbf{w} - \mathbf{u}\}$

b. $\{\mathbf{u} + \mathbf{v}, \mathbf{v} + \mathbf{w}, \mathbf{w} + \mathbf{u}\}$

c. $\{\mathbf{u} - \mathbf{v}, \mathbf{v} - \mathbf{w}, \mathbf{w} - \mathbf{z}, \mathbf{z} - \mathbf{u}\}$

d. $\{\mathbf{u} + \mathbf{v}, \mathbf{v} + \mathbf{w}, \mathbf{w} + \mathbf{z}, \mathbf{z} + \mathbf{u}\}$

Exercise 6.3.24 Let U and W be subspaces of V with bases $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ and $\{\mathbf{w}_1, \mathbf{w}_2\}$ respectively. If U and W have only the zero vector in common, show that $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{w}_1, \mathbf{w}_2\}$ is independent.

Exercise 6.3.25 Let $\{p, q\}$ be independent polynomials. Show that $\{p, q, pq\}$ is independent if and only if $\deg p \geq 1$ and $\deg q \geq 1$.

Exercise 6.3.26 If z is a complex number, show that $\{z, z^2\}$ is independent if and only if z is not real.

Exercise 6.3.27 Let $B = \{A_1, A_2, \dots, A_n\} \subseteq \mathbf{M}_{mn}$, and write $B' = \{A_1^T, A_2^T, \dots, A_n^T\} \subseteq \mathbf{M}_{nm}$. Show that:

a. B is independent if and only if B' is independent.

b. B spans $\mathbf{M}_{mn}$ if and only if B' spans $\mathbf{M}_{nm}$.

Exercise 6.3.28 If $V = \mathbf{F}[a, b]$ as in Example 6.1.7, show that the set of constant functions is a subspace of dimension 1 (f is **constant** if there is a number c such that $f(x) = c$ for all x).

Exercise 6.3.29

a. If U is an invertible $n \times n$ matrix and $\{A_1, A_2, \dots, A_{mn}\}$ is a basis of $\mathbf{M}_{mn}$, show that $\{A_1U, A_2U, \dots, A_{mn}U\}$ is also a basis.

b. Show that part (a) fails if U is not invertible. [*Hint*: Theorem 2.4.5.]

Exercise 6.3.30 Show that $\{(a, b), (a_1, b_1)\}$ is a basis of $\mathbb{R}^2$ if and only if $\{a + bx, a_1 + b_1x\}$ is a basis of $\mathbf{P}_1$.

Exercise 6.3.31 Find the dimension of the subspace $\text{span}\{1, \sin^2\theta, \cos 2\theta\}$ of $\mathbf{F}[0, 2\pi]$.

Exercise 6.3.32 Show that $\mathbf{F}[0, 1]$ is not finite dimensional.

Exercise 6.3.33 If U and W are subspaces of V, define their intersection $U \cap W$ as follows:

$$U \cap W = \{\mathbf{v} \mid \mathbf{v} \text{ is in both } U \text{ and } W\}$$

a. Show that $U \cap W$ is a subspace contained in U and W.

b. Show that $U \cap W = \{\mathbf{0}\}$ if and only if $\{\mathbf{u},\ \mathbf{w}\}$ is independent for any nonzero vectors $\mathbf{u}$ in U and $\mathbf{w}$ in W.

c. If B and D are bases of U and W, and if $U \cap W = \{\mathbf{0}\}$, show that $B \cup D = \{\mathbf{v} \mid \mathbf{v} \text{ is in } B \text{ or } D\}$ is independent.

Exercise 6.3.34 If U and W are vector spaces, let $V = \{(\mathbf{u},\ \mathbf{w}) \mid \mathbf{u} \text{ in } U \text{ and } \mathbf{w} \text{ in } W\}$.

a. Show that V is a vector space if $(\mathbf{u},\ \mathbf{w}) + (\mathbf{u}_1,\ \mathbf{w}_1) = (\mathbf{u} + \mathbf{u}_1,\ \mathbf{w} + \mathbf{w}_1)$ and $a(\mathbf{u},\ \mathbf{w}) = (a\mathbf{u},\ a\mathbf{w})$.

b. If $\dim U = m$ and $\dim W = n$, show that $\dim V = m + n$.

c. If $V_1,\ \dots,\ V_m$ are vector spaces, let

$$\begin{aligned} V &= V_1 \times \cdots \times V_m \\ &= \{(\mathbf{v}_1,\ \dots,\ \mathbf{v}_m) \mid \mathbf{v}_i \in V_i \text{ for each } i\} \end{aligned}$$

denote the space of n-tuples from the V_i with componentwise operations (see Exercise 6.1.17). If $\dim V_i = n_i$ for each i, show that $\dim V = n_1 + \cdots + n_m$.

Exercise 6.3.35 Let $\mathbf{D}_n$ denote the set of all functions f from the set $\{1,\ 2,\ \dots,\ n\}$ to $\mathbb{R}$.

a. Show that $\mathbf{D}_n$ is a vector space with pointwise addition and scalar multiplication.

b. Show that $\{S_1,\ S_2,\ \dots,\ S_n\}$ is a basis of $\mathbf{D}_n$ where, for each $k = 1,\ 2,\ \dots,\ n$, the function S_k is defined by $S_k(k) = 1$, whereas $S_k(j) = 0$ if $j \neq k$.

Exercise 6.3.36 A polynomial $p(x)$ is called **even** if $p(-x) = p(x)$ and **odd** if $p(-x) = -p(x)$. Let E_n and O_n denote the sets of even and odd polynomials in $\mathbf{P}_n$.

a. Show that E_n is a subspace of $\mathbf{P}_n$ and find $\dim E_n$.

b. Show that O_n is a subspace of $\mathbf{P}_n$ and find $\dim O_n$.

Exercise 6.3.37 Let $\{\mathbf{v}_1,\ \dots,\ \mathbf{v}_n\}$ be independent in a vector space V, and let A be an $n \times n$ matrix. Define $\mathbf{u}_1,\ \dots,\ \mathbf{u}_n$ by

$$\left[\begin{array}{c} \mathbf{u}_1 \\ \vdots \\ \mathbf{u}_n \end{array}\right] = A \left[\begin{array}{c} \mathbf{v}_1 \\ \vdots \\ \mathbf{v}_n \end{array}\right]$$

(See Exercise 6.1.18.) Show that $\{\mathbf{u}_1,\ \dots,\ \mathbf{u}_n\}$ is independent if and only if A is invertible.

6.4 Finite Dimensional Spaces

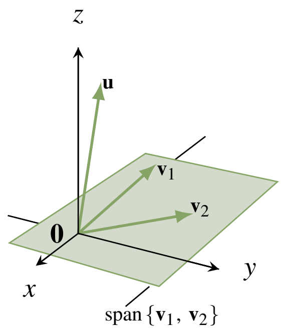
Up to this point, we have had no guarantee that an arbitrary vector space *has* a basis—and hence no guarantee that one can speak *at all* of the dimension of V. However, Theorem 6.4.1 will show that any space that is spanned by a finite set of vectors has a (finite) basis: The proof requires the following basic lemma, of interest in itself, that gives a way to enlarge a given independent set of vectors.

Lemma 6.4.1: Independent Lemma

Let $\{\mathbf{v}_1,\ \mathbf{v}_2,\ \dots,\ \mathbf{v}_k\}$ be an independent set of vectors in a vector space V. If $\mathbf{u} \in V$ but[5] *$\mathbf{u} \notin \text{span}\{\mathbf{v}_1,\ \mathbf{v}_2,\ \dots,\ \mathbf{v}_k\}$, then $\{\mathbf{u},\ \mathbf{v}_1,\ \mathbf{v}_2,\ \dots,\ \mathbf{v}_k\}$ is also independent.*

Proof. Let $t\mathbf{u} + t_1\mathbf{v}_1 + t_2\mathbf{v}_2 + \cdots + t_k\mathbf{v}_k = \mathbf{0}$; we must show that all the coefficients are zero. First, $t = 0$ because, otherwise, $\mathbf{u} = -\frac{t_1}{t}\mathbf{v}_1 - \frac{t_2}{t}\mathbf{v}_2 - \cdots - \frac{t_k}{t}\mathbf{v}_k$ is in $\text{span}\{\mathbf{v}_1,\ \mathbf{v}_2,\ \dots,\ \mathbf{v}_k\}$, contrary to our assumption. Hence $t = 0$. But then $t_1\mathbf{v}_1 + t_2\mathbf{v}_2 + \cdots + t_k\mathbf{v}_k = \mathbf{0}$ so the rest of the t_i are zero by the independence of $\{\mathbf{v}_1,\ \mathbf{v}_2,\ \dots,\ \mathbf{v}_k\}$. This is what we wanted. □

[5] If X is a set, we write $a \in X$ to indicate that a is an element of the set X. If a is not an element of X, we write $a \notin X$.



Note that the converse of Lemma 6.4.1 is also true: if $\{\mathbf{u}, \mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ is independent, then $\mathbf{u}$ is not in span $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$.

As an illustration, suppose that $\{\mathbf{v}_1, \mathbf{v}_2\}$ is independent in $\mathbb{R}^3$. Then $\mathbf{v}_1$ and $\mathbf{v}_2$ are not parallel, so span $\{\mathbf{v}_1, \mathbf{v}_2\}$ is a plane through the origin (shaded in the diagram). By Lemma 6.4.1, $\mathbf{u}$ is not in this plane if and only if $\{\mathbf{u}, \mathbf{v}_1, \mathbf{v}_2\}$ is independent.

Definition 6.7 Finite Dimensional and Infinite Dimensional Vector Spaces

A vector space V is called ***finite dimensional*** *if it is spanned by a finite set of vectors. Otherwise, V is called* ***infinite dimensional****.*

Thus the zero vector space $\{\mathbf{0}\}$ is finite dimensional because $\{\mathbf{0}\}$ is a spanning set.

Lemma 6.4.2

Let V be a finite dimensional vector space. If U is any subspace of V, then any independent subset of U can be enlarged to a finite basis of U.

Proof. Suppose that I is an independent subset of U. If span $I = U$ then I is already a basis of U. If span $I \neq U$, choose $\mathbf{u}_1 \in U$ such that $\mathbf{u}_1 \notin$ span I. Hence the set $I \cup \{\mathbf{u}_1\}$ is independent by Lemma 6.4.1. If span $(I \cup \{\mathbf{u}_1\}) = U$ we are done; otherwise choose $\mathbf{u}_2 \in U$ such that $\mathbf{u}_2 \notin$ span $(I \cup \{\mathbf{u}_1\})$. Hence $I \cup \{\mathbf{u}_1, \mathbf{u}_2\}$ is independent, and the process continues. We claim that a basis of U will be reached eventually. Indeed, if no basis of U is ever reached, the process creates arbitrarily large independent sets in V. But this is impossible by the fundamental theorem because V is finite dimensional and so is spanned by a finite set of vectors. □

Theorem 6.4.1

Let V be a finite dimensional vector space spanned by m vectors.

1. *V has a finite basis, and* $\dim V \leq m$.
2. *Every independent set of vectors in V can be enlarged to a basis of V by adding vectors from any fixed basis of V.*
3. *If U is a subspace of V, then*
   a. *U is finite dimensional and* $\dim U \leq \dim V$.
   b. *If* $\dim U = \dim V$ *then* $U = V$.

Proof.

1. If $V = \{\mathbf{0}\}$, then V has an empty basis and $\dim V = 0 \leq m$. Otherwise, let $\mathbf{v} \neq \mathbf{0}$ be a vector in V. Then $\{\mathbf{v}\}$ is independent, so (1) follows from Lemma 6.4.2 with $U = V$.

2. We refine the proof of Lemma 6.4.2. Fix a basis B of V and let I be an independent subset of V. If $\text{span}\, I = V$ then I is already a basis of V. If $\text{span}\, I \neq V$, then B is not contained in I (because B spans V). Hence choose $\mathbf{b}_1 \in B$ such that $\mathbf{b}_1 \notin \text{span}\, I$. Hence the set $I \cup \{\mathbf{b}_1\}$ is independent by Lemma 6.4.1. If $\text{span}\,(I \cup \{\mathbf{b}_1\}) = V$ we are done; otherwise a similar argument shows that $(I \cup \{\mathbf{b}_1,\ \mathbf{b}_2\})$ is independent for some $\mathbf{b}_2 \in B$. Continue this process. As in the proof of Lemma 6.4.2, a basis of V will be reached eventually.

3. a. This is clear if $U = \{\mathbf{0}\}$. Otherwise, let $\mathbf{u} \neq \mathbf{0}$ in U. Then $\{\mathbf{u}\}$ can be enlarged to a finite basis B of U by Lemma 6.4.2, proving that U is finite dimensional. But B is independent in V, so $\dim U \leq \dim V$ by the fundamental theorem.

   b. This is clear if $U = \{\mathbf{0}\}$ because V *has* a basis; otherwise, it follows from (2). □

Theorem 6.4.1 shows that a vector space V is finite dimensional if and only if it has a finite basis (possibly empty), and that every subspace of a finite dimensional space is again finite dimensional.

Example 6.4.1

Enlarge the independent set $D = \left\{ \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \right\}$ to a basis of $\mathbf{M}_{22}$.

Solution. The standard basis of $\mathbf{M}_{22}$ is $\left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$, so including one of these in D will produce a basis by Theorem 6.4.1. In fact including *any* of these matrices in D produces an independent set (verify), and hence a basis by Theorem 6.4.4. Of course these vectors are not the only possibilities, for example, including $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ works as well.

Example 6.4.2

Find a basis of $\mathbf{P}_3$ containing the independent set $\{1 + x,\ 1 + x^2\}$.

Solution. The standard basis of $\mathbf{P}_3$ is $\{1,\ x,\ x^2,\ x^3\}$, so including two of these vectors will do. If we use 1 and x^3, the result is $\{1,\ 1 + x,\ 1 + x^2,\ x^3\}$. This is independent because the polynomials have distinct degrees (Example 6.3.4), and so is a basis by Theorem 6.4.1. Of course, including $\{1,\ x\}$ or $\{1,\ x^2\}$ would *not* work!

Example 6.4.3

Show that the space $\mathbf{P}$ of all polynomials is infinite dimensional.

Solution. For each $n \geq 1$, $\mathbf{P}$ has a subspace $\mathbf{P}_n$ of dimension $n + 1$. Suppose $\mathbf{P}$ is finite dimensional, say $\dim \mathbf{P} = m$. Then $\dim \mathbf{P}_n \leq \dim \mathbf{P}$ by Theorem 6.4.1, that is $n+1 \leq m$. This is impossible since n is arbitrary, so $\mathbf{P}$ must be infinite dimensional.

The next example illustrates how (2) of Theorem 6.4.1 can be used.

Example 6.4.4

If $\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_k$ are independent columns in $\mathbb{R}^n$, show that they are the first k columns in some invertible $n \times n$ matrix.

Solution. By Theorem 6.4.1, expand $\{\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_k\}$ to a basis $\{\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_k, \mathbf{c}_{k+1}, \dots, \mathbf{c}_n\}$ of $\mathbb{R}^n$. Then the matrix $A = \left[\begin{array}{cccccc} \mathbf{c}_1 & \mathbf{c}_2 & \dots & \mathbf{c}_k & \mathbf{c}_{k+1} & \dots & \mathbf{c}_n \end{array} \right]$ with this basis as its columns is an $n \times n$ matrix and it is invertible by Theorem 5.2.3.

Theorem 6.4.2

Let U and W be subspaces of the finite dimensional space V.

1. *If $U \subseteq W$, then $\dim U \leq \dim W$.*
2. *If $U \subseteq W$ and $\dim U = \dim W$, then $U = W$.*

Proof. Since W is finite dimensional, (1) follows by taking $V = W$ in part (3) of Theorem 6.4.1. Now assume $\dim U = \dim W = n$, and let B be a basis of U. Then B is an independent set in W. If $U \neq W$, then $\operatorname{span} B \neq W$, so B can be extended to an independent set of $n+1$ vectors in W by Lemma 6.4.1. This contradicts the fundamental theorem (Theorem 6.3.2) because W is spanned by $\dim W = n$ vectors. Hence $U = W$, proving (2). □

Theorem 6.4.2 is very useful. This was illustrated in Example 5.2.13 for $\mathbb{R}^2$ and $\mathbb{R}^3$; here is another example.

Example 6.4.5

If a is a number, let W denote the subspace of all polynomials in $\mathbf{P}_n$ that have a as a root:

$$W = \{p(x) \mid p(x) \in \mathbf{P}_n \text{ and } p(a) = 0\}$$

Show that $\{(x-a), (x-a)^2, \dots, (x-a)^n\}$ is a basis of W.

Solution. Observe first that $(x-a), (x-a)^2, \dots, (x-a)^n$ are members of W, and that they are independent because they have distinct degrees (Example 6.3.4). Write

$$U = \operatorname{span}\{(x-a), (x-a)^2, \dots, (x-a)^n\}$$

Then we have $U \subseteq W \subseteq \mathbf{P}_n$, $\dim U = n$, and $\dim \mathbf{P}_n = n+1$. Hence $n \leq \dim W \leq n+1$ by Theorem 6.4.2. Since $\dim W$ is an integer, we must have $\dim W = n$ or $\dim W = n+1$. But then $W = U$ or $W = \mathbf{P}_n$, again by Theorem 6.4.2. Because $W \neq \mathbf{P}_n$, it follows that $W = U$, as required.

A set of vectors is called **dependent** if it is *not* independent, that is if some nontrivial linear combination vanishes. The next result is a convenient test for dependence.

Lemma 6.4.3: Dependent Lemma

A set $D = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ of vectors in a vector space V is dependent if and only if some vector in D is a linear combination of the others.

Proof. Let $\mathbf{v}_2$ (say) be a linear combination of the rest: $\mathbf{v}_2 = s_1\mathbf{v}_1 + s_3\mathbf{v}_3 + \cdots + s_k\mathbf{v}_k$. Then

$$s_1\mathbf{v}_1 + (-1)\mathbf{v}_2 + s_3\mathbf{v}_3 + \cdots + s_k\mathbf{v}_k = \mathbf{0}$$

is a nontrivial linear combination that vanishes, so D is dependent. Conversely, if D is dependent, let $t_1\mathbf{v}_1 + t_2\mathbf{v}_2 + \cdots + t_k\mathbf{v}_k = \mathbf{0}$ where some coefficient is nonzero. If (say) $t_2 \neq 0$, then $\mathbf{v}_2 = -\frac{t_1}{t_2}\mathbf{v}_1 - \frac{t_3}{t_2}\mathbf{v}_3 - \cdots - \frac{t_k}{t_2}\mathbf{v}_k$ is a linear combination of the others. □

Lemma 6.4.1 gives a way to enlarge independent sets to a basis; by contrast, Lemma 6.4.3 shows that spanning sets can be cut down to a basis.

Theorem 6.4.3

Let V be a finite dimensional vector space. Any spanning set for V can be cut down (by deleting vectors) to a basis of V.

Proof. Since V is finite dimensional, it has a finite spanning set S. Among all spanning sets contained in S, choose S_0 containing the smallest number of vectors. It suffices to show that S_0 is independent (then S_0 is a basis, proving the theorem). Suppose, on the contrary, that S_0 is not independent. Then, by Lemma 6.4.3, some vector $\mathbf{u} \in S_0$ is a linear combination of the set $S_1 = S_0 \setminus \{\mathbf{u}\}$ of vectors in S_0 other than $\mathbf{u}$. It follows that $\text{span } S_0 = \text{span } S_1$, that is, $V = \text{span } S_1$. But S_1 has fewer elements than S_0 so this contradicts the choice of S_0. Hence S_0 is independent after all. □

Note that, with Theorem 6.4.1, Theorem 6.4.3 completes the promised proof of Theorem 5.2.6 for the case $V = \mathbb{R}^n$.

Example 6.4.6

Find a basis of $\mathbf{P}_3$ in the spanning set $S = \{1,\ x+x^2,\ 2x-3x^2,\ 1+3x-2x^2,\ x^3\}$.

Solution. Since $\dim \mathbf{P}_3 = 4$, we must eliminate one polynomial from S. It cannot be x^3 because the span of the rest of S is contained in $\mathbf{P}_2$. But eliminating $1+3x-2x^2$ does leave a basis (verify). Note that $1+3x-2x^2$ is the sum of the first three polynomials in S.

Theorems 6.4.1 and 6.4.3 have other useful consequences.

Theorem 6.4.4

Let V be a vector space with $\dim V = n$, and suppose S is a set of exactly n vectors in V. Then S is independent if and only if S spans V.

Proof. Assume first that S is independent. By Theorem 6.4.1, S is contained in a basis B of V. Hence $|S| = n = |B|$ so, since $S \subseteq B$, it follows that $S = B$. In particular S spans V.

Conversely, assume that S spans V, so S contains a basis B by Theorem 6.4.3. Again $|S| = n = |B|$ so, since $S \supseteq B$, it follows that $S = B$. Hence S is independent. □

One of independence or spanning is often easier to establish than the other when showing that a set of vectors is a basis. For example if $V = \mathbb{R}^n$ it is easy to check whether a subset S of $\mathbb{R}^n$ is orthogonal (hence independent) but checking spanning can be tedious. Here are three more examples.

Example 6.4.7

Consider the set $S = \{p_0(x),\ p_1(x),\ \dots,\ p_n(x)\}$ of polynomials in $\mathbf{P}_n$. If $\deg p_k(x) = k$ for each k, show that S is a basis of $\mathbf{P}_n$.

Solution. The set S is independent—the degrees are distinct—see Example 6.3.4. Hence S is a basis of $\mathbf{P}_n$ by Theorem 6.4.4 because $\dim \mathbf{P}_n = n+1$.

Example 6.4.8

Let V denote the space of all symmetric 2×2 matrices. Find a basis of V consisting of invertible matrices.

Solution. We know that $\dim V = 3$ (Example 6.3.11), so what is needed is a set of three invertible, symmetric matrices that (using Theorem 6.4.4) is either independent or spans V. The set $\left\{ \left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array} \right], \left[\begin{array}{cc} 1 & 0 \\ 0 & -1 \end{array} \right], \left[\begin{array}{cc} 0 & 1 \\ 1 & 0 \end{array} \right] \right\}$ is independent (verify) and so is a basis of the required type.

Example 6.4.9

Let A be any $n \times n$ matrix. Show that there exist n^2+1 scalars $a_0,\ a_1,\ a_2,\ \dots,\ a_{n^2}$ not all zero, such that

$$a_0 I + a_1 A + a_2 A^2 + \cdots + a_{n^2} A^{n^2} = 0$$

where I denotes the $n \times n$ identity matrix.

Solution. The space $\mathbf{M}_{nn}$ of all $n \times n$ matrices has dimension n^2 by Example 6.3.7. Hence the n^2+1 matrices $I,\ A,\ A^2,\ \dots,\ A^{n^2}$ cannot be independent by Theorem 6.4.4, so a nontrivial linear combination vanishes. This is the desired conclusion.

The result in Example 6.4.9 can be written as $f(A) = 0$ where $f(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_{n^2} x^{n^2}$. In other words, A satisfies a nonzero polynomial $f(x)$ of degree at most n^2. In fact we know that A satisfies a nonzero polynomial of degree n (this is the Cayley-Hamilton theorem—see Theorem 8.7.10), but the brevity of the solution in Example 6.4.6 is an indication of the power of these methods.

If U and W are subspaces of a vector space V, there are two related subspaces that are of interest, their **sum** $U+W$ and their **intersection** $U \cap W$, defined by

$$U + W = \{\mathbf{u} + \mathbf{w} \mid \mathbf{u} \in U \text{ and } \mathbf{w} \in W\}$$

$$U \cap W = \{\mathbf{v} \in V \mid \mathbf{v} \in U \text{ and } \mathbf{v} \in W\}$$

It is routine to verify that these are indeed subspaces of V, that $U \cap W$ is contained in both U and W, and that $U + W$ contains both U and W. We conclude this section with a useful fact about the dimensions of these spaces. The proof is a good illustration of how the theorems in this section are used.

Theorem 6.4.5

Suppose that U and W are finite dimensional subspaces of a vector space V. Then $U+W$ is finite dimensional and

$$\dim(U+W) = \dim U + \dim W - \dim(U \cap W).$$

Proof. Since $U \cap W \subseteq U$, it has a finite basis, say $\{\mathbf{x}_1, \dots, \mathbf{x}_d\}$. Extend it to a basis $\{\mathbf{x}_1, \dots, \mathbf{x}_d, \mathbf{u}_1, \dots, \mathbf{u}_m\}$ of U by Theorem 6.4.1. Similarly extend $\{\mathbf{x}_1, \dots, \mathbf{x}_d\}$ to a basis $\{\mathbf{x}_1, \dots, \mathbf{x}_d, \mathbf{w}_1, \dots, \mathbf{w}_p\}$ of W. Then

$$U + W = \operatorname{span}\{\mathbf{x}_1, \dots, \mathbf{x}_d, \mathbf{u}_1, \dots, \mathbf{u}_m, \mathbf{w}_1, \dots, \mathbf{w}_p\}$$

as the reader can verify, so $U+W$ is finite dimensional. For the rest, it suffices to show that $\{\mathbf{x}_1, \dots, \mathbf{x}_d, \mathbf{u}_1, \dots, \mathbf{u}_m, \mathbf{w}_1, \dots, \mathbf{w}_p\}$ is independent (verify). Suppose that

$$r_1\mathbf{x}_1 + \cdots + r_d\mathbf{x}_d + s_1\mathbf{u}_1 + \cdots + s_m\mathbf{u}_m + t_1\mathbf{w}_1 + \cdots + t_p\mathbf{w}_p = \mathbf{0} \tag{6.1}$$

where the r_i, s_j, and t_k are scalars. Then

$$r_1\mathbf{x}_1 + \cdots + r_d\mathbf{x}_d + s_1\mathbf{u}_1 + \cdots + s_m\mathbf{u}_m = -(t_1\mathbf{w}_1 + \cdots + t_p\mathbf{w}_p)$$

is in U (left side) and also in W (right side), and so is in $U \cap W$. Hence $(t_1\mathbf{w}_1 + \cdots + t_p\mathbf{w}_p)$ is a linear combination of $\{\mathbf{x}_1, \dots, \mathbf{x}_d\}$, so $t_1 = \cdots = t_p = 0$, because $\{\mathbf{x}_1, \dots, \mathbf{x}_d, \mathbf{w}_1, \dots, \mathbf{w}_p\}$ is independent. Similarly, $s_1 = \cdots = s_m = 0$, so (6.1) becomes $r_1\mathbf{x}_1 + \cdots + r_d\mathbf{x}_d = \mathbf{0}$. It follows that $r_1 = \cdots = r_d = 0$, as required. □

Theorem 6.4.5 is particularly interesting if $U \cap W = \{\mathbf{0}\}$. Then there are *no* vectors $\mathbf{x}_i$ in the above proof, and the argument shows that if $\{\mathbf{u}_1, \dots, \mathbf{u}_m\}$ and $\{\mathbf{w}_1, \dots, \mathbf{w}_p\}$ are bases of U and W respectively, then $\{\mathbf{u}_1, \dots, \mathbf{u}_m, \mathbf{w}_1, \dots, \mathbf{w}_p\}$ is a basis of $U + W$. In this case $U+W$ is said to be a **direct sum** (written $U \oplus W$); we return to this in Chapter 9.

Exercises for 6.4

Exercise 6.4.1 In each case, find a basis for V that includes the vector $\mathbf{v}$.

a. $V = \mathbb{R}^3$, $\mathbf{v} = (1, -1, 1)$

b. $V = \mathbb{R}^3$, $\mathbf{v} = (0, 1, 1)$

c. $V = \mathbf{M}_{22}$, $\mathbf{v} = \left[\begin{array}{cc} 1 & 1 \\ 1 & 1 \end{array}\right]$

d. $V = \mathbf{P}_2$, $\mathbf{v} = x^2 - x + 1$

Exercise 6.4.2 In each case, find a basis for V among the given vectors.

a. $V = \mathbb{R}^3$,
$\{(1, 1, -1), (2, 0, 1), (-1, 1, -2), (1, 2, 1)\}$

b. $V = \mathbf{P}_2$, $\{x^2 + 3, x + 2, x^2 - 2x - 1, x^2 + x\}$

Exercise 6.4.3 In each case, find a basis of V containing $\mathbf{v}$ and $\mathbf{w}$.

a. $V = \mathbb{R}^4$, $\mathbf{v} = (1, -1, 1, -1)$, $\mathbf{w} = (0, 1, 0, 1)$

b. $V = \mathbb{R}^4$, $\mathbf{v} = (0, 0, 1, 1)$, $\mathbf{w} = (1, 1, 1, 1)$

c. $V = \mathbf{M}_{22}$, $\mathbf{v} = \left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array}\right]$, $\mathbf{w} = \left[\begin{array}{cc} 0 & 1 \\ 1 & 0 \end{array}\right]$

d. $V = \mathbf{P}_3$, $\mathbf{v} = x^2 + 1$, $\mathbf{w} = x^2 + x$

Exercise 6.4.4

a. If z is not a real number, show that $\{z, z^2\}$ is a basis of the real vector space $\mathbb{C}$ of all complex numbers.

b. If z is neither real nor pure imaginary, show that $\{z, \overline{z}\}$ is a basis of $\mathbb{C}$.

Exercise 6.4.5 In each case use Theorem 6.4.4 to decide if S is a basis of V.

a. $V = \mathbf{M}_{22}$;
$S = \left\{ \left[\begin{array}{cc} 1 & 1 \\ 1 & 1 \end{array}\right], \left[\begin{array}{cc} 0 & 1 \\ 1 & 1 \end{array}\right], \left[\begin{array}{cc} 0 & 0 \\ 1 & 1 \end{array}\right], \left[\begin{array}{cc} 0 & 0 \\ 0 & 1 \end{array}\right] \right\}$

b. $V = \mathbf{P}_3$; $S = \{2x^2, 1+x, 3, 1+x+x^2+x^3\}$

Exercise 6.4.6

a. Find a basis of $\mathbf{M}_{22}$ consisting of matrices with the property that $A^2 = A$.

b. Find a basis of $\mathbf{P}_3$ consisting of polynomials whose coefficients sum to 4. What if they sum to 0?

Exercise 6.4.7 If $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ is a basis of V, determine which of the following are bases.

a. $\{\mathbf{u}+\mathbf{v}, \mathbf{u}+\mathbf{w}, \mathbf{v}+\mathbf{w}\}$

b. $\{2\mathbf{u}+\mathbf{v}+3\mathbf{w}, 3\mathbf{u}+\mathbf{v}-\mathbf{w}, \mathbf{u}-4\mathbf{w}\}$

c. $\{\mathbf{u}, \mathbf{u}+\mathbf{v}+\mathbf{w}\}$

d. $\{\mathbf{u}, \mathbf{u}+\mathbf{w}, \mathbf{u}-\mathbf{w}, \mathbf{v}+\mathbf{w}\}$

Exercise 6.4.8

a. Can two vectors span $\mathbb{R}^3$? Can they be linearly independent? Explain.

b. Can four vectors span $\mathbb{R}^3$? Can they be linearly independent? Explain.

Exercise 6.4.9 Show that any nonzero vector in a finite dimensional vector space is part of a basis.

Exercise 6.4.10 If A is a square matrix, show that $\det A = 0$ if and only if some row is a linear combination of the others.

Exercise 6.4.11 Let D, I, and X denote finite, nonempty sets of vectors in a vector space V. Assume that D is dependent and I is independent. In each case answer yes or no, and defend your answer.

a. If $X \supseteq D$, must X be dependent?

b. If $X \subseteq D$, must X be dependent?

c. If $X \supseteq I$, must X be independent?

d. If $X \subseteq I$, must X be independent?

Exercise 6.4.12 If U and W are subspaces of V and $\dim U = 2$, show that either $U \subseteq W$ or $\dim(U \cap W) \leq 1$.

Exercise 6.4.13 Let A be a nonzero 2×2 matrix and write $U = \{X \text{ in } \mathbf{M}_{22} \mid XA = AX\}$. Show that $\dim U \geq 2$. [*Hint*: I and A are in U.]

Exercise 6.4.14 If $U \subseteq \mathbb{R}^2$ is a subspace, show that $U = \{\mathbf{0}\}$, $U = \mathbb{R}^2$, or U is a line through the origin.

Exercise 6.4.15 Given $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \dots, \mathbf{v}_k$, and $\mathbf{v}$, let $U = \text{span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ and $W = \text{span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k, \mathbf{v}\}$. Show that either $\dim W = \dim U$ or $\dim W = 1 + \dim U$.

Exercise 6.4.16 Suppose U is a subspace of $\mathbf{P}_1$, $U \neq \{0\}$, and $U \neq \mathbf{P}_1$. Show that either $U = \mathbb{R}$ or $U = \mathbb{R}(a+x)$ for some a in $\mathbb{R}$.

Exercise 6.4.17 Let U be a subspace of V and assume $\dim V = 4$ and $\dim U = 2$. Does every basis of V result from adding (two) vectors to some basis of U? Defend your answer.

Exercise 6.4.18 Let U and W be subspaces of a vector space V.

a. If $\dim V = 3$, $\dim U = \dim W = 2$, and $U \neq W$, show that $\dim(U \cap W) = 1$.

b. Interpret (a.) geometrically if $V = \mathbb{R}^3$.

Exercise 6.4.19 Let $U \subseteq W$ be subspaces of V with $\dim U = k$ and $\dim W = m$, where $k < m$. If $k < l < m$, show that a subspace X exists where $U \subseteq X \subseteq W$ and $\dim X = l$.

Exercise 6.4.20 Let $B = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ be a *maximal* independent set in a vector space V. That is, no set of more than n vectors S is independent. Show that B is a basis of V.

Exercise 6.4.21 Let $B = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ be a *minimal* spanning set for a vector space V. That is, V cannot be spanned by fewer than n vectors. Show that B is a basis of V.

Exercise 6.4.22

a. Let $p(x)$ and $q(x)$ lie in $\mathbf{P}_1$ and suppose that $p(1) \neq 0$, $q(2) \neq 0$, and $p(2) = 0 = q(1)$. Show that $\{p(x), q(x)\}$ is a basis of $\mathbf{P}_1$. [*Hint*: If $rp(x) + sq(x) = 0$, evaluate at $x = 1$, $x = 2$.]

b. Let $B = \{p_0(x), p_1(x), \dots, p_n(x)\}$ be a set of polynomials in $\mathbf{P}_n$. Assume that there exist numbers $a_0, a_1, \dots, a_n$ such that $p_i(a_i) \neq 0$ for each i but $p_i(a_j) = 0$ if i is different from j. Show that B is a basis of $\mathbf{P}_n$.

Exercise 6.4.23 Let V be the set of all infinite sequences $(a_0, a_1, a_2, \dots)$ of real numbers. Define addition and scalar multiplication by

$$(a_0, a_1, \dots) + (b_0, b_1, \dots) = (a_0 + b_0, a_1 + b_1, \dots)$$

and

$$r(a_0, a_1, \dots) = (ra_0, ra_1, \dots)$$

a. Show that V is a vector space.

b. Show that V is not finite dimensional.

c. [For those with some calculus.] Show that the set of convergent sequences (that is, $\lim\limits_{n\to\infty} a_n$ exists) is a subspace, also of infinite dimension.

Exercise 6.4.24 Let A be an $n \times n$ matrix of rank r. If $U = \{X \text{ in } \mathbf{M}_{nn} \mid AX = 0\}$, show that $\dim U = n(n-r)$. [*Hint*: Exercise 6.3.34.]

Exercise 6.4.25 Let U and W be subspaces of V.

a. Show that $U + W$ is a subspace of V containing both U and W.

b. Show that $\text{span}\,\{\mathbf{u}, \mathbf{w}\} = \mathbb{R}\mathbf{u} + \mathbb{R}\mathbf{w}$ for any vectors $\mathbf{u}$ and $\mathbf{w}$.

c. Show that

$$\begin{aligned}&\text{span}\,\{\mathbf{u}_1, \dots, \mathbf{u}_m, \mathbf{w}_1, \dots, \mathbf{w}_n\}\\ &= \text{span}\,\{\mathbf{u}_1, \dots, \mathbf{u}_m\} + \text{span}\,\{\mathbf{w}_1, \dots, \mathbf{w}_n\}\end{aligned}$$

for any vectors $\mathbf{u}_i$ in U and $\mathbf{w}_j$ in W.

Exercise 6.4.26 If A and B are $m \times n$ matrices, show that $\text{rank}\,(A + B) \leq \text{rank}\,A + \text{rank}\,B$. [*Hint*: If U and V are the column spaces of A and B, respectively, show that the column space of $A + B$ is contained in $U + V$ and that $\dim\,(U + V) \leq \dim U + \dim V$. (See Theorem 6.4.5.)]

6.5 An Application to Polynomials

The vector space of all polynomials of degree at most n is denoted $\mathbf{P}_n$, and it was established in Section 6.3 that $\mathbf{P}_n$ has dimension $n + 1$; in fact, $\{1, x, x^2, \dots, x^n\}$ is a basis. More generally, *any* $n + 1$ polynomials of distinct degrees form a basis, by Theorem 6.4.4 (they are independent by Example 6.3.4). This proves

Theorem 6.5.1

Let $p_0(x), p_1(x), p_2(x), \dots, p_n(x)$ be polynomials in $\mathbf{P}_n$ of degrees $0, 1, 2, \dots, n$, respectively. Then $\{p_0(x), \dots, p_n(x)\}$ is a basis of $\mathbf{P}_n$.

An immediate consequence is that $\{1, (x - a), (x - a)^2, \dots, (x - a)^n\}$ is a basis of $\mathbf{P}_n$ for any number a. Hence we have the following:

Corollary 6.5.1

If a is any number, every polynomial $f(x)$ of degree at most n has an expansion in powers of $(x-a)$:

$$f(x) = a_0 + a_1(x-a) + a_2(x-a)^2 + \cdots + a_n(x-a)^n \tag{6.2}$$

If $f(x)$ is evaluated at $x = a$, then equation (6.2) becomes

$$f(x) = a_0 + a_1(a-a) + \cdots + a_n(a-a)^n = a_0$$

Hence $a_0 = f(a)$, and equation (6.2) can be written $f(x) = f(a) + (x-a)g(x)$, where $g(x)$ is a polynomial of degree $n-1$ (this assumes that $n \geq 1$). If it happens that $f(a) = 0$, then it is clear that $f(x)$ has the form $f(x) = (x-a)g(x)$. Conversely, every such polynomial certainly satisfies $f(a) = 0$, and we obtain:

Corollary 6.5.2

Let $f(x)$ be a polynomial of degree $n \geq 1$ and let a be any number. Then:

Remainder Theorem

1. *$f(x) = f(a) + (x-a)g(x)$ for some polynomial $g(x)$ of degree $n-1$.*

Factor Theorem

2. *$f(a) = 0$ if and only if $f(x) = (x-a)g(x)$ for some polynomial $g(x)$.*

The polynomial $g(x)$ can be computed easily by using "long division" to divide $f(x)$ by $(x-a)$—see Appendix D.

All the coefficients in the expansion (6.2) of $f(x)$ in powers of $(x-a)$ can be determined in terms of the derivatives of $f(x)$.[6] These will be familiar to students of calculus. Let $f^{(n)}(x)$ denote the nth derivative of the polynomial $f(x)$, and write $f^{(0)}(x) = f(x)$. Then, if

$$f(x) = a_0 + a_1(x-a) + a_2(x-a)^2 + \cdots + a_n(x-a)^n$$

it is clear that $a_0 = f(a) = f^{(0)}(a)$. Differentiation gives

$$f^{(1)}(x) = a_1 + 2a_2(x-a) + 3a_3(x-a)^2 + \cdots + na_n(x-a)^{n-1}$$

and substituting $x = a$ yields $a_1 = f^{(1)}(a)$. This continues to give $a_2 = \frac{f^{(2)}(a)}{2!}$, $a_3 = \frac{f^{(3)}(a)}{3!}$, $\dots$, $a_k = \frac{f^{(k)}(a)}{k!}$, where $k!$ is defined as $k! = k(k-1)\cdots 2 \cdot 1$. Hence we obtain the following:

Corollary 6.5.3: Taylor's Theorem

If $f(x)$ is a polynomial of degree n, then

$$f(x) = f(a) + \frac{f^{(1)}(a)}{1!}(x-a) + \frac{f^{(2)}(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n$$

[6] The discussion of Taylor's theorem can be omitted with no loss of continuity.

Example 6.5.1

Expand $f(x) = 5x^3 + 10x + 2$ as a polynomial in powers of $x - 1$.

Solution. The derivatives are $f^{(1)}(x) = 15x^2 + 10$, $f^{(2)}(x) = 30x$, and $f^{(3)}(x) = 30$. Hence the Taylor expansion is

$$\begin{aligned} f(x) &= f(1) + \frac{f^{(1)}(1)}{1!}(x-1) + \frac{f^{(2)}(1)}{2!}(x-1)^2 + \frac{f^{(3)}(1)}{3!}(x-1)^3 \\ &= 17 + 25(x-1) + 15(x-1)^2 + 5(x-1)^3 \end{aligned}$$

Taylor's theorem is useful in that it provides a formula for the coefficients in the expansion. It is dealt with in calculus texts and will not be pursued here.

Theorem 6.5.1 produces bases of $\mathbf{P}_n$ consisting of polynomials of distinct degrees. A different criterion is involved in the next theorem.

Theorem 6.5.2

Let $f_0(x),\ f_1(x),\ \dots,\ f_n(x)$ be nonzero polynomials in $\mathbf{P}_n$. Assume that numbers $a_0,\ a_1,\ \dots,\ a_n$ exist such that

$$\begin{aligned} f_i(a_i) &\neq 0 \quad \text{for each } i \\ f_i(a_j) &= 0 \quad \text{if } i \neq j \end{aligned}$$

Then

1. $\{f_0(x),\ \dots,\ f_n(x)\}$ *is a basis of* $\mathbf{P}_n$.
2. *If $f(x)$ is any polynomial in $\mathbf{P}_n$, its expansion as a linear combination of these basis vectors is*

$$f(x) = \frac{f(a_0)}{f_0(a_0)}f_0(x) + \frac{f(a_1)}{f_1(a_1)}f_1(x) + \cdots + \frac{f(a_n)}{f_n(a_n)}f_n(x)$$

Proof.

1. It suffices (by Theorem 6.4.4) to show that $\{f_0(x),\ \dots,\ f_n(x)\}$ is linearly independent (because $\dim \mathbf{P}_n = n+1$). Suppose that

$$r_0f_0(x) + r_1f_1(x) + \cdots + r_nf_n(x) = 0,\ r_i \in \mathbb{R}$$

Because $f_i(a_0) = 0$ for all $i > 0$, taking $x = a_0$ gives $r_0f_0(a_0) = 0$. But then $r_0 = 0$ because $f_0(a_0) \neq 0$. The proof that $r_i = 0$ for $i > 0$ is analogous.

2. By (1), $f(x) = r_0f_0(x) + \cdots + r_nf_n(x)$ for *some* numbers r_i. Once again, evaluating at a_0 gives $f(a_0) = r_0f_0(a_0)$, so $r_0 = f(a_0)/f_0(a_0)$. Similarly, $r_i = f(a_i)/f_i(a_i)$ for each i. □

Example 6.5.2

Show that $\{x^2 - x,\ x^2 - 2x,\ x^2 - 3x + 2\}$ is a basis of $\mathbf{P}_2$.

Solution. Write $f_0(x) = x^2 - x = x(x-1)$, $f_1(x) = x^2 - 2x = x(x-2)$, and $f_2(x) = x^2 - 3x + 2 = (x-1)(x-2)$. Then the conditions of Theorem 6.5.2 are satisfied with $a_0 = 2$, $a_1 = 1$, and $a_2 = 0$.

We investigate one natural choice of the polynomials $f_i(x)$ in Theorem 6.5.2. To illustrate, let a_0, a_1, and a_2 be distinct numbers and write

$$f_0(x) = \frac{(x-a_1)(x-a_2)}{(a_0-a_1)(a_0-a_2)} \quad f_1(x) = \frac{(x-a_0)(x-a_2)}{(a_1-a_0)(a_1-a_2)} \quad f_2(x) = \frac{(x-a_0)(x-a_1)}{(a_2-a_0)(a_2-a_1)}$$

Then $f_0(a_0) = f_1(a_1) = f_2(a_2) = 1$, and $f_i(a_j) = 0$ for $i \neq j$. Hence Theorem 6.5.2 applies, and because $f_i(a_i) = 1$ for each i, the formula for expanding any polynomial is simplified.

In fact, this can be generalized with no extra effort. If $a_0, a_1, \dots, a_n$ are distinct numbers, define the **Lagrange polynomials** $\delta_0(x), \delta_1(x), \dots, \delta_n(x)$ relative to these numbers as follows:

$$\delta_k(x) = \frac{\prod_{i \neq k}(x-a_i)}{\prod_{i \neq k}(a_k-a_i)} \quad k = 0, 1, 2, \dots, n$$

Here the numerator is the product of all the terms $(x-a_0), (x-a_1), \dots, (x-a_n)$ with $(x-a_k)$ omitted, and a similar remark applies to the denominator. If $n = 2$, these are just the polynomials in the preceding paragraph. For another example, if $n = 3$, the polynomial $\delta_1(x)$ takes the form

$$\delta_1(x) = \frac{(x-a_0)(x-a_2)(x-a_3)}{(a_1-a_0)(a_1-a_2)(a_1-a_3)}$$

In the general case, it is clear that $\delta_i(a_i) = 1$ for each i and that $\delta_i(a_j) = 0$ if $i \neq j$. Hence Theorem 6.5.2 specializes as Theorem 6.5.3.

Theorem 6.5.3: Lagrange Interpolation Expansion

Let $a_0, a_1, \dots, a_n$ be distinct numbers. The corresponding set

$$\{\delta_0(x), \delta_1(x), \dots, \delta_n(x)\}$$

of Lagrange polynomials is a basis of $\mathbf{P}_n$, and any polynomial $f(x)$ in $\mathbf{P}_n$ has the following unique expansion as a linear combination of these polynomials.

$$f(x) = f(a_0)\delta_0(x) + f(a_1)\delta_1(x) + \cdots + f(a_n)\delta_n(x)$$

Example 6.5.3

Find the Lagrange interpolation expansion for $f(x) = x^2 - 2x + 1$ relative to $a_0 = -1$, $a_1 = 0$, and $a_2 = 1$.

Solution. The Lagrange polynomials are

$$\begin{aligned}\delta_0 &= \frac{(x-0)(x-1)}{(-1-0)(-1-1)} = \frac{1}{2}(x^2-x)\\ \delta_1 &= \frac{(x+1)(x-1)}{(0+1)(0-1)} = -(x^2-1)\\ \delta_2 &= \frac{(x+1)(x-0)}{(1+1)(1-0)} = \frac{1}{2}(x^2+x)\end{aligned}$$

Because $f(-1)=4$, $f(0)=1$, and $f(1)=0$, the expansion is

$$f(x) = 2(x^2-x)-(x^2-1)$$

The Lagrange interpolation expansion gives an easy proof of the following important fact.

Theorem 6.5.4

Let $f(x)$ be a polynomial in $\mathbf{P}_n$, and let $a_0, a_1, \dots, a_n$ denote distinct numbers. If $f(a_i)=0$ for all i, then $f(x)$ is the zero polynomial (that is, all coefficients are zero).

Proof. All the coefficients in the Lagrange expansion of $f(x)$ are zero. □

Exercises for 6.5

Exercise 6.5.1 If polynomials $f(x)$ and $g(x)$ satisfy $f(a)=g(a)$, show that $f(x)-g(x)=(x-a)h(x)$ for some polynomial $h(x)$.

Exercises 6.5.2, 6.5.3, 6.5.4, and 6.5.5 require polynomial differentiation.

Exercise 6.5.2 Expand each of the following as a polynomial in powers of $x-1$.

a. $f(x)=x^3-2x^2+x-1$

b. $f(x)=x^3+x+1$

c. $f(x)=x^4$

d. $f(x)=x^3-3x^2+3x$

Exercise 6.5.3 Prove Taylor's theorem for polynomials.

Exercise 6.5.4 Use Taylor's theorem to derive the **binomial theorem**:

$$(1+x)^n = \binom{n}{0}+\binom{n}{1}x+\binom{n}{2}x^2+\cdots+\binom{n}{n}x^n$$

Here the **binomial coefficients** $\binom{n}{r}$ are defined by

$$\binom{n}{r} = \frac{n!}{r!(n-r)!}$$

where $n! = n(n-1)\cdots 2\cdot 1$ if $n \geq 1$ and $0! = 1$.

Exercise 6.5.5 Let $f(x)$ be a polynomial of degree n. Show that, given any polynomial $g(x)$ in $\mathbf{P}_n$, there exist numbers $b_0, b_1, \dots, b_n$ such that

$$g(x) = b_0 f(x) + b_1 f^{(1)}(x) + \cdots + b_n f^{(n)}(x)$$

where $f^{(k)}(x)$ denotes the kth derivative of $f(x)$.

Exercise 6.5.6 Use Theorem 6.5.2 to show that the following are bases of $\mathbf{P}_2$.

a. $\{x^2-2x,\ x^2+2x,\ x^2-4\}$

b. $\{x^2-3x+2,\ x^2-4x+3,\ x^2-5x+6\}$

Exercise 6.5.7 Find the Lagrange interpolation expansion of $f(x)$ relative to $a_0=1$, $a_1=2$, and $a_2=3$ if:

a. $f(x)=x^2+1$

b. $f(x)=x^2+x+1$

Exercise 6.5.8 Let $a_0, a_1, \dots, a_n$ be distinct numbers. If $f(x)$ and $g(x)$ in $\mathbf{P}_n$ satisfy $f(a_i) = g(a_i)$ for all i, show that $f(x) = g(x)$. [*Hint*: See Theorem 6.5.4.]

Exercise 6.5.9 Let $a_0, a_1, \dots, a_n$ be distinct numbers. If $f(x) \in \mathbf{P}_{n+1}$ satisfies $f(a_i) = 0$ for each $i = 0, 1, \dots, n$, show that $f(x) = r(x-a_0)(x-a_1)\cdots(x-a_n)$ for some r in $\mathbb{R}$. [*Hint*: r is the coefficient of x^{n+1} in $f(x)$. Consider $f(x) - r(x-a_0)\cdots(x-a_n)$ and use Theorem 6.5.4.]

Exercise 6.5.10 Let a and b denote distinct numbers.

a. Show that $\{(x-a), (x-b)\}$ is a basis of $\mathbf{P}_1$.

b. Show that $\{(x-a)^2, (x-a)(x-b), (x-b)^2\}$ is a basis of $\mathbf{P}_2$.

c. Show that $\{(x-a)^n, (x-a)^{n-1}(x-b), \dots, (x-a)(x-b)^{n-1}, (x-b)^n\}$ is a basis of $\mathbf{P}_n$. [*Hint*: If a linear combination vanishes, evaluate at $x = a$ and $x = b$. Then reduce to the case $n-2$ by using the fact that if $p(x)q(x) = 0$ in $\mathbf{P}$, then either $p(x) = 0$ or $q(x) = 0$.]

Exercise 6.5.11 Let a and b be two distinct numbers. Assume that $n \geq 2$ and let

$$U_n = \{f(x) \text{ in } \mathbf{P}_n \mid f(a) = 0 = f(b)\}.$$

a. Show that

$$U_n = \{(x-a)(x-b)p(x) \mid p(x) \text{ in } \mathbf{P}_{n-2}\}$$

b. Show that $\dim U_n = n-1$.

[*Hint*: If $p(x)q(x) = 0$ in $\mathbf{P}$, then either $p(x) = 0$, or $q(x) = 0$.]

c. Show $\{(x-a)^{n-1}(x-b), (x-a)^{n-2}(x-b)^2, \dots, (x-a)^2(x-b)^{n-2}, (x-a)(x-b)^{n-1}\}$ is a basis of U_n. [*Hint*: Exercise 6.5.10.]

6.6 An Application to Differential Equations

Call a function $f : \mathbb{R} \to \mathbb{R}$ **differentiable** if it can be differentiated as many times as we want. If f is a differentiable function, the nth derivative $f^{(n)}$ of f is the result of differentiating n times. Thus $f^{(0)} = f$, $f^{(1)} = f'$, $f^{(2)} = f^{(1)\prime}$, $\dots$ and, in general, $f^{(n+1)} = f^{(n)\prime}$ for each $n \geq 0$. For small values of n these are often written as $f, f', f'', f''', \dots$.

If a, b, and c are numbers, the differential equations

$$f'' + af' + bf = 0 \quad \text{or} \quad f''' + af'' + bf' + cf = 0$$

are said to be of **second-order** and **third-order**, respectively. In general, an equation

$$f^{(n)} + a_{n-1}f^{(n-1)} + a_{n-2}f^{(n-2)} + \cdots + a_2f^{(2)} + a_1f^{(1)} + a_0f^{(0)} = 0, \quad a_i \text{ in } \mathbb{R} \tag{6.3}$$

is called a **differential equation of order** n. In this section we investigate the set of solutions to (6.3) and, if n is 1 or 2, find explicit solutions. Of course an acquaintance with calculus is required.

Let f and g be solutions to (6.3). Then $f+g$ is also a solution because $(f+g)^{(k)} = f^{(k)} + g^{(k)}$ for all k, and af is a solution for any a in $\mathbb{R}$ because $(af)^{(k)} = af^{(k)}$. It follows that the set of solutions to (6.3) is a vector space, and we ask for the dimension of this space.

We have already dealt with the simplest case (see Theorem 3.5.1):

Theorem 6.6.1

The set of solutions of the first-order differential equation $f' + af = 0$ is a one-dimensional vector space and $\{e^{-ax}\}$ is a basis.

There is a far-reaching generalization of Theorem 6.6.1 that will be proved in Theorem 7.4.1.

Theorem 6.6.2

The set of solutions to the nth order equation (6.3) has dimension n.

Remark

Every differential equation of order n can be converted into a system of n linear first-order equations (see Exercises 3.5.6 and 3.5.7). In the case that the matrix of this system is diagonalizable, this approach provides a proof of Theorem 6.6.2. But if the matrix is not diagonalizable, Theorem 7.4.1 is required.

Theorem 6.6.1 suggests that we look for solutions to (6.3) of the form $e^{\lambda x}$ for some number λ. This is a good idea. If we write $f(x) = e^{\lambda x}$, it is easy to verify that $f^{(k)}(x) = \lambda^k e^{\lambda x}$ for each $k \geq 0$, so substituting f in (6.3) gives

$$(\lambda^n + a_{n-1}\lambda^{n-1} + a_{n-2}\lambda^{n-2} + \cdots + a_2\lambda^2 + a_1\lambda^1 + a_0)e^{\lambda x} = 0$$

Since $e^{\lambda x} \neq 0$ for all x, this shows that $e^{\lambda x}$ is a solution of (6.3) if and only if λ is a root of the **characteristic polynomial** $c(x)$, defined to be

$$c(x) = x^n + a_{n-1}x^{n-1} + a_{n-2}x^{n-2} + \cdots + a_2x^2 + a_1x + a_0$$

This proves Theorem 6.6.3.

Theorem 6.6.3

If λ is real, the function $e^{\lambda x}$ is a solution of (6.3) if and only if λ is a root of the characteristic polynomial $c(x)$.

Example 6.6.1

Find a basis of the space U of solutions of $f''' - 2f'' - f' - 2f = 0$.

Solution. The characteristic polynomial is $x^3 - 2x^2 - x - 1 = (x-1)(x+1)(x-2)$, with roots $\lambda_1 = 1$, $\lambda_2 = -1$, and $\lambda_3 = 2$. Hence e^x, e^{-x}, and e^{2x} are all in U. Moreover they are independent (by Lemma 6.6.1 below) so, since $\dim(U) = 3$ by Theorem 6.6.2, $\{e^x, e^{-x}, e^{2x}\}$ is a basis of U.

Lemma 6.6.1

If $\lambda_1, \lambda_2, \dots, \lambda_k$ are distinct, then $\{e^{\lambda_1 x}, e^{\lambda_2 x}, \dots, e^{\lambda_k x}\}$ is linearly independent.

Proof. If $r_1e^{\lambda_1 x} + r_2e^{\lambda_2 x} + \cdots + r_ke^{\lambda_k x} = 0$ for all x, then $r_1 + r_2e^{(\lambda_2-\lambda_1)x} + \cdots + r_ke^{(\lambda_k-\lambda_1)x} = 0$; that is, $r_2e^{(\lambda_2-\lambda_1)x} + \cdots + r_ke^{(\lambda_k-\lambda_1)x}$ is a constant. Since the λ_i are distinct, this forces $r_2 = \cdots = r_k = 0$, whence $r_1 = 0$ also. This is what we wanted. □

Theorem 6.6.4

Let U denote the space of solutions to the second-order equation

$$f'' + af' + bf = 0$$

where a and b are real constants. Assume that the characteristic polynomial $x^2 + ax + b$ has two real roots λ and μ. Then

1. *If $\lambda \neq \mu$, then $\{e^{\lambda x},\ e^{\mu x}\}$ is a basis of U.*
2. *If $\lambda = \mu$, then $\{e^{\lambda x},\ xe^{\lambda x}\}$ is a basis of U.*

Proof. Since $\dim(U) = 2$ by Theorem 6.6.2, (1) follows by Lemma 6.6.1, and (2) follows because the set $\{e^{\lambda x},\ xe^{\lambda x}\}$ is independent (Exercise 6.6.3). □

Example 6.6.2

Find the solution of $f'' + 4f' + 4f = 0$ that satisfies the **boundary conditions** $f(0) = 1$, $f(1) = -1$.

Solution. The characteristic polynomial is $x^2 + 4x + 4 = (x+2)^2$, so -2 is a double root. Hence $\{e^{-2x},\ xe^{-2x}\}$ is a basis for the space of solutions, and the general solution takes the form $f(x) = ce^{-2x} + dxe^{-2x}$. Applying the boundary conditions gives $1 = f(0) = c$ and $-1 = f(1) = (c+d)e^{-2}$. Hence $c = 1$ and $d = -(1+e^2)$, so the required solution is

$$f(x) = e^{-2x} - (1+e^2)xe^{-2x}$$

One other question remains: What happens if the roots of the characteristic polynomial are not real? To answer this, we must first state precisely what $e^{\lambda x}$ means when λ is not real. If q is a real number, define

$$e^{iq} = \cos q + i \sin q$$

where $i^2 = -1$. Then the relationship $e^{iq}e^{iq_1} = e^{i(q+q_1)}$ holds for all real q and q_1, as is easily verified. If $\lambda = p + iq$, where p and q are real numbers, we define

$$e^{\lambda} = e^{p}e^{iq} = e^{p}(\cos q + i \sin q)$$

Then it is a routine exercise to show that

1. $e^{\lambda}e^{\mu} = e^{\lambda+\mu}$
2. $e^{\lambda} = 1$ if and only if $\lambda = 0$
3. $(e^{\lambda x})' = \lambda e^{\lambda x}$

These easily imply that $f(x) = e^{\lambda x}$ is a solution to $f'' + af' + bf = 0$ if λ is a (possibly complex) root of the characteristic polynomial $x^2 + ax + b$. Now write $\lambda = p + iq$ so that

$$f(x) = e^{\lambda x} = e^{px}\cos(qx) + ie^{px}\sin(qx)$$

For convenience, denote the real and imaginary parts of $f(x)$ as $u(x) = e^{px}\cos(qx)$ and $v(x) = e^{px}\sin(qx)$. Then the fact that $f(x)$ satisfies the differential equation gives

$$0 = f'' + af' + bf = (u'' + au' + bu) + i(v'' + av' + bv)$$

Equating real and imaginary parts shows that $u(x)$ and $v(x)$ are both solutions to the differential equation. This proves part of Theorem 6.6.5.

Theorem 6.6.5

Let U denote the space of solutions of the second-order differential equation

$$f'' + af' + bf = 0$$

where a and b are real. Suppose λ is a nonreal root of the characteristic polynomial $x^2 + ax + b$. If $\lambda = p + iq$, where p and q are real, then

$$\{e^{px}\cos(qx),\ e^{px}\sin(qx)\}$$

is a basis of U.

Proof. The foregoing discussion shows that these functions lie in U. Because $\dim U = 2$ by Theorem 6.6.2, it suffices to show that they are linearly independent. But if

$$re^{px}\cos(qx) + se^{px}\sin(qx) = 0$$

for all x, then $r\cos(qx) + s\sin(qx) = 0$ for all x (because $e^{px} \neq 0$). Taking $x = 0$ gives $r = 0$, and taking $x = \frac{\pi}{2q}$ gives $s = 0$ ($q \neq 0$ because λ is not real). This is what we wanted. □

Example 6.6.3

Find the solution $f(x)$ to $f'' - 2f' + 2f = 0$ that satisfies $f(0) = 2$ and $f(\frac{\pi}{2}) = 0$.

Solution. The characteristic polynomial $x^2 - 2x + 2$ has roots $1 + i$ and $1 - i$. Taking $\lambda = 1 + i$ (quite arbitrarily) gives $p = q = 1$ in the notation of Theorem 6.6.5, so $\{e^x\cos x,\ e^x\sin x\}$ is a basis for the space of solutions. The general solution is thus $f(x) = e^x(r\cos x + s\sin x)$. The boundary conditions yield $2 = f(0) = r$ and $0 = f(\frac{\pi}{2}) = e^{\pi/2}s$. Thus $r = 2$ and $s = 0$, and the required solution is $f(x) = 2e^x\cos x$.

The following theorem is an important special case of Theorem 6.6.5.

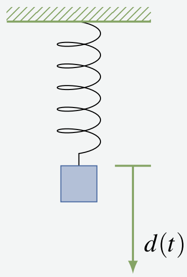
Theorem 6.6.6

If $q \neq 0$ is a real number, the space of solutions to the differential equation $f'' + q^2 f = 0$ has basis $\{\cos(qx),\ \sin(qx)\}$.

Proof. The characteristic polynomial $x^2 + q^2$ has roots qi and $-qi$, so Theorem 6.6.5 applies with $p = 0$. □

In many situations, the displacement $s(t)$ of some object at time t turns out to have an oscillating form $s(t) = c\sin(at) + d\cos(at)$. These are called **simple harmonic motions**. An example follows.

Example 6.6.4



A weight is attached to an extension spring (see diagram). If it is pulled from the equilibrium position and released, it is observed to oscillate up and down. Let $d(t)$ denote the distance of the weight below the equilibrium position t seconds later. It is known (**Hooke's law**) that the acceleration $d''(t)$ of the weight is proportional to the displacement $d(t)$ and in the opposite direction. That is,

$$d''(t) = -kd(t)$$

where $k > 0$ is called the **spring constant**. Find $d(t)$ if the maximum extension is 10 cm below the equilibrium position and find the **period** of the oscillation (time taken for the weight to make a full oscillation).

Solution. It follows from Theorem 6.6.6 (with $q^2 = k$) that

$$d(t) = r\sin(\sqrt{k}\,t) + s\cos(\sqrt{k}\,t)$$

where r and s are constants. The condition $d(0) = 0$ gives $s = 0$, so $d(t) = r\sin(\sqrt{k}\,t)$. Now the maximum value of the function $\sin x$ is 1 (when $x = \frac{\pi}{2}$), so $r = 10$ (when $t = \frac{\pi}{2\sqrt{k}}$). Hence

$$d(t) = 10\sin(\sqrt{k}\,t)$$

Finally, the weight goes through a full oscillation as $\sqrt{k}\,t$ increases from 0 to 2π. The time taken is $t = \frac{2\pi}{\sqrt{k}}$, the period of the oscillation.

Exercises for 6.6

Exercise 6.6.1 Find a solution f to each of the following differential equations satisfying the given boundary conditions.

a. $f' - 3f = 0$; $f(1) = 2$

b. $f' + f = 0$; $f(1) = 1$

c. $f'' + 2f' - 15f = 0$; $f(1) = f(0) = 0$

d. $f'' + f' - 6f = 0$; $f(0) = 0$, $f(1) = 1$

e. $f'' - 2f' + f = 0$; $f(1) = f(0) = 1$

f. $f'' - 4f' + 4f = 0$; $f(0) = 2$, $f(-1) = 0$

g. $f'' - 3af' + 2a^2 f = 0$; $a \neq 0$; $f(0) = 0$, $f(1) = 1 - e^a$

h. $f'' - a^2 f = 0$, $a \neq 0$; $f(0) = 1$, $f(1) = 0$

i. $f'' - 2f' + 5f = 0$; $f(0) = 1$, $f(\frac{\pi}{4}) = 0$

j. $f'' + 4f' + 5f = 0$; $f(0) = 0$, $f(\frac{\pi}{2}) = 1$

Exercise 6.6.2 If the characteristic polynomial of $f'' + af' + bf = 0$ has real roots, show that $f = 0$ is the only solution satisfying $f(0) = 0 = f(1)$.

Exercise 6.6.3 Complete the proof of Theorem 6.6.2. [*Hint*: If λ is a double root of $x^2 + ax + b$, show that $a = -2\lambda$ and $b = \lambda^2$. Hence $xe^{\lambda x}$ is a solution.]

Exercise 6.6.4

a. Given the equation $f' + af = b$, $(a \neq 0)$, make the substitution $f(x) = g(x) + b/a$ and obtain a differential equation for g. Then derive the general solution for $f' + af = b$.

b. Find the general solution to $f' + f = 2$.

Exercise 6.6.5 Consider the differential equation $f' + af' + bf = g$, where g is some fixed function. Assume that f_0 is one solution of this equation.

a. Show that the general solution is $cf_1 + df_2 + f_0$, where c and d are constants and $\{f_1, f_2\}$ is any basis for the solutions to $f'' + af' + bf = 0$.

b. Find a solution to $f'' + f' - 6f = 2x^3 - x^2 - 2x$. [*Hint*: Try $f(x) = \frac{-1}{3}x^3$.]

Exercise 6.6.6 A radioactive element decays at a rate proportional to the amount present. Suppose an initial mass of 10 grams decays to 8 grams in 3 hours.

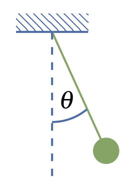
a. Find the mass t hours later.

b. Find the *half-life* of the element—the time it takes to decay to half its mass.

Exercise 6.6.7 The population $N(t)$ of a region at time t increases at a rate proportional to the population. If the population doubles in 5 years and is 3 million initially, find $N(t)$.

Exercise 6.6.8 Consider a spring, as in Example 6.6.4. If the period of the oscillation is 30 seconds, find the spring constant k.

Exercise 6.6.9 As a pendulum swings (see the diagram), let t measure the time since it was vertical. The angle $\theta = \theta(t)$ from the vertical can be shown to satisfy the equation $\theta'' + k\theta = 0$, provided that θ is small. If the maximal angle is $\theta = 0.05$ radians, find $\theta(t)$ in terms of k. If the period is 0.5 seconds, find k. [Assume that $\theta = 0$ when $t = 0$.]



Supplementary Exercises for Chapter 6

Exercise 6.1 (Requires calculus) Let V denote the space of all functions $f : \mathbb{R} \to \mathbb{R}$ for which the derivatives f' and f'' exist. Show that f_1, f_2, and f_3 in V are linearly independent provided that their **wronskian** $w(x)$ is nonzero for some x, where

$$w(x) = \det \begin{bmatrix} f_1(x) & f_2(x) & f_3(x) \\ f_1'(x) & f_2'(x) & f_3'(x) \\ f_1''(x) & f_2''(x) & f_3''(x) \end{bmatrix}$$

Exercise 6.2 Let $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ be a basis of $\mathbb{R}^n$ (written as columns), and let A be an $n \times n$ matrix.

a. If A is invertible, show that $\{A\mathbf{v}_1, A\mathbf{v}_2, \dots, A\mathbf{v}_n\}$ is a basis of $\mathbb{R}^n$.

b. If $\{A\mathbf{v}_1, A\mathbf{v}_2, \dots, A\mathbf{v}_n\}$ is a basis of $\mathbb{R}^n$, show that A is invertible.

Exercise 6.3 If A is an $m \times n$ matrix, show that A has rank m if and only if col A contains every column of I_m.

Exercise 6.4 Show that null A = null (A^TA) for any real matrix A.

Exercise 6.5 Let A be an $m \times n$ matrix of rank r. Show that $\dim(\text{null } A) = n - r$ (Theorem 5.4.3) as follows. Choose a basis $\{\mathbf{x}_1, \dots, \mathbf{x}_k\}$ of null A and extend it to a basis $\{\mathbf{x}_1, \dots, \mathbf{x}_k, \mathbf{z}_1, \dots, \mathbf{z}_m\}$ of $\mathbb{R}^n$. Show that $\{A\mathbf{z}_1, \dots, A\mathbf{z}_m\}$ is a basis of col A.