CHAPTER 1

Vector Spaces

Introduction

In this first chapter of *A Course in Linear Algebra*, we begin by introducing the fundamental concept of a vector space, a mathematical structure that has proved to be very useful in describing some of the common features of important mathematical objects such as the set of vectors in the plane and the set of all functions from the real line to itself.

Our first goal will be to write down a list of properties that hold for the algebraic sum and scalar multiplication operations in the two aforementioned examples. We will then take this list of properties as our definition of what a general vector space should be. This is a typical example of the idea of defining an object by specifying what properties it should have, a commonly used notion in mathematics.

We will then develop a repertoire of examples of vector spaces, drawing on ideas from geometry and calculus. Following this, we will explore the inner structure of vector spaces by studying subspaces and spanning sets and bases (special subsets from which the whole vector space can be built up). Along the way, we will find that most of the calculations that we need to perform involve solving simultaneous systems of linear equations, so we will also discuss a general method for doing this.

The concept of a vector space provides a way to organize, explain, and build on many topics you have seen before in geometry, algebra, and calculus. At the same time, as you begin to study linear algebra, you may find that the way everything is presented seems very general and abstract. Of course, to the extent that this is true, it is a reflection of the fact that mathematicians have seen a very basic general pattern that holds in many different situations. They have exploited this information by inventing the ideas discussed in this chapter in order to understand all these situations and treat them all without resorting to dealing with each case separately. With time and practice, working in general vector spaces should become natural to you, just as the equally abstract concept of number (as opposed to specific collections of some number of objects) has become second nature.

§1.1. VECTOR SPACES

The basic geometric objects that are studied in linear algebra are called vector spaces. Since you have probably seen vectors before in your mathematical experience, we begin by recalling some basic facts about vectors in the plane to help motivate the discussion of general vector spaces that follows.

In the geometry of the Euclidean plane, a *vector* is usually defined as a directed line segment or "arrow," that is, as a line segment with one endpoint distinguished as the "head" or final point, and the other distinguished as the "tail" or initial point. See Figure 1.1. Vectors are useful for describing quantities with both a magnitude and a direction. Geometrically, the *length* of the directed line segment may be taken to represent the magnitude of the quantity; the direction is given by the direction that the arrow is pointing. Important examples of quantities of this kind are the instantaneous velocity and the instantaneous acceleration at each time of an object moving along a path in the plane, the momentum of the moving object, forces, and so on. In physics these quantities are treated mathematically by using vectors as just described.

In linear algebra one of our major concerns will be the *algebraic properties* of vectors. By this we mean, for example, the operations by which vectors may be combined to produce new vectors and the properties of those operations. For instance, if we consider the set of *all* vectors in the plane with a tail at some fixed point O, then it is possible to combine vectors to produce new vectors in two ways.

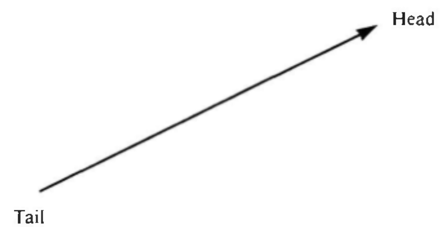

Figure 1.1

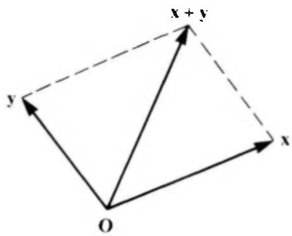

Figure 1.2

First, if we take two vectors **x** and **y**, then we can define their *vector sum* **x** + **y** to be the vector whose tail is at the point O and whose head is at the fourth corner of the parallelogram with sides **x** and **y**. See Figure 1.2. One physical interpretation of this sum operation is as follows. If two forces, represented by vectors **x** and **y**, act on an object located at the point O then the resulting force will be given by the vector sum **x** + **y**.

Second, if we take a vector **x** and a positive real number c (called a *scalar* in this context), then we can define the *product* of the vector **x** and the scalar c to be the vector in the same direction as **x** but with a magnitude or length that is equal to c times the magnitude of **x**. If $c > 1$, this has the effect of magnifying **x**, whereas if $c < 1$, this shrinks **x**. The case $c > 1$ is pictured in Figure 1.3. Physically, a positive scalar multiple of a vector may be thought of in the following way. For example, in the case $c = 2$, if the vector **x** represents a force, then the vector 2**x** represents a force that is "twice as strong" and that acts in the same direction. Similarly, the vector (1/2)**x** represents a force that is "one-half as strong." The product c**x** may also be defined if $c < 0$. In this case the vector c**x** will point along the same line through the origin as **x** but in the opposite direction from **x**. The magnitude of c**x** in this case will be equal to $|c|$ times the magnitude of **x**. See Figure 1.4.

Further properties of these two operations on vectors may be derived directly from these geometric definitions. However, to bring their algebraic nature into clearer focus, we will now consider an alternate way to understand these operations. If we introduce the familiar Cartesian coordinate system in the plane and place the origin at the point O = (0, 0), then a vector whose tail is at O is uniquely specified

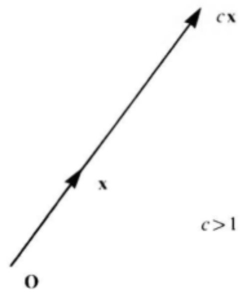

Figure 1.3

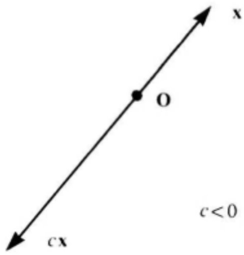

Figure 1.4

by the coordinates of its head. That is, vectors may be described as ordered pairs of real numbers. See Figure 1.5.

In this way we obtain a one-to-one correspondence between the set of vectors with a tail at the origin and the set $\mathbf{R}^2$ (the set of ordered pairs of real numbers) and we write $\mathbf{x} = (x_1, x_2)$ to indicate the vector whose head is at the point (x_1, x_2).

Our two operations on vectors may be described using coordinates. First, from the parallelogram law for the vector sum, we see that if $\mathbf{x} = (x_1, x_2)$ and $\mathbf{y} = (y_1, y_2)$, then $\mathbf{x} + \mathbf{y} = (x_1 + y_1, x_2 + y_2)$. See Figure 1.6. That is, to find the vector sum, we simply add "component-wise." For example the vector sum $(2, 5) + (4, -3)$ is equal to $(6, 2)$. Second, if c is a scalar and $\mathbf{x} = (x_1, x_2)$ is a vector, then $c\mathbf{x} = (cx_1, cx_2)$. The scalar multiple $4(-1, 2)$ is equal to $(-4, 8)$.

With this description, the familiar properties of addition and multiplication of real numbers may be used to show that our two operations on vectors have the following algebraic properties:

1. The vector sum is *associative:* For all $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{z} \in \mathbf{R}^2$ we have

$$(\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z})$$

2. The vector sum is *commutative:* For all $\mathbf{x}$ and $\mathbf{y} \in \mathbf{R}^2$ we have

$$\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$$

3. There is an *additive identity element* $\mathbf{0} = (0, 0) \in \mathbf{R}^2$ with the property that for all $\mathbf{x} \in \mathbf{R}^2$, $\mathbf{x} + \mathbf{0} = \mathbf{x}$.

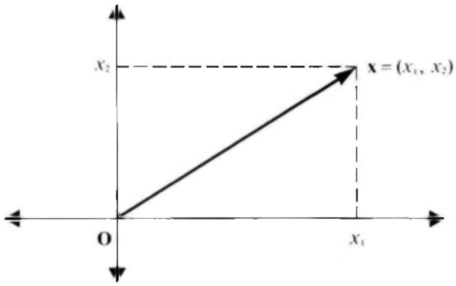

Figure 1.5

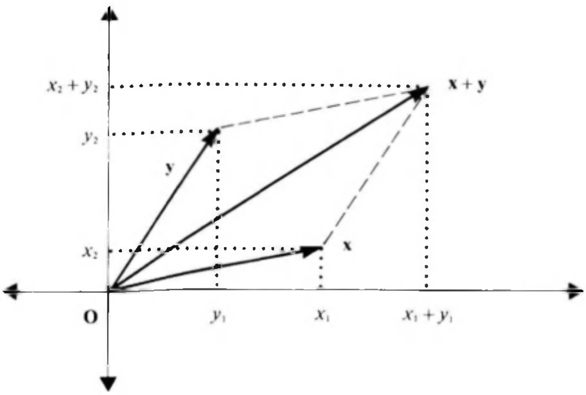

Figure 1.6

4. For each $\mathbf{x} = (x_1, x_2) \in \mathbf{R}^2$ there is an *additive inverse* $-\mathbf{x} = (-x_1, -x_2)$ with the property that $\mathbf{x} + (-\mathbf{x}) = \mathbf{0}$.

5. Multiplication by a scalar is *distributive* over vector sums: For all $c \in \mathbf{R}$ and all $\mathbf{x}, \mathbf{y} \in \mathbf{R}^2$ we have

$$c(\mathbf{x} + \mathbf{y}) = c\mathbf{x} + c\mathbf{y}$$

6. Multiplication of a vector by a sum of scalars is also *distributive:* For all $c, d \in \mathbf{R}$ and all $\mathbf{x} \in \mathbf{R}^2$ we have

$$(c + d)\mathbf{x} = c\mathbf{x} + d\mathbf{x}$$

7. For all $c, d \in \mathbf{R}$ and all $\mathbf{x} \in \mathbf{R}^2$ we have

$$c(d\mathbf{x}) = (cd)\mathbf{x}$$

8. For all $\mathbf{x} \in \mathbf{R}^2$ we have $1\mathbf{x} = \mathbf{x}$.

For example, property 1 follows from the fact that addition of real numbers is also associative. If $\mathbf{x} = (x_1, x_2)$, $\mathbf{y} = (y_1, y_2)$, and $\mathbf{z} = (z_1, z_2)$, then we have

$$\begin{aligned}
(\mathbf{x} + \mathbf{y}) + \mathbf{z} &= ((x_1, x_2) + (y_1, y_2)) + (z_1, z_2) \\
&= ((x_1 + y_1) + z_1, (x_2 + y_2) + z_2) \\
&= (x_1 + (y_1 + z_1), x_2 + (y_2 + z_2)) \\
&= \mathbf{x} + (\mathbf{y} + \mathbf{z})
\end{aligned}$$

Property 5 follows from the fact that multiplication is distributive over addition in $\mathbf{R}$. If $\mathbf{x} = (x_1, x_2)$, $\mathbf{y} = (y_1, y_2)$, and c is a scalar, then

$$\begin{aligned}
c(\mathbf{x} + \mathbf{y}) &= c((x_1, x_2) + (y_1, y_2)) \\
&= c(x_1 + y_1, x_2 + y_2) \\
&= (c(x_1 + y_1), c(x_2 + y_2))
\end{aligned}$$

$$= (cx_1 + cy_1, cx_2 + cy_2) \qquad \text{(by distributivity)}$$

$$= (cx_1, cx_2) + (cy_1, cy_2)$$

$$= c\mathbf{x} + c\mathbf{y}$$

The reader is encouraged to verify that the other aforementioned properties also hold in general.

These eight properties describe the algebraic behavior of vectors in the plane completely. Although the list was derived from this specific example, it is interesting that many other important mathematical structures share the same characteristic properties. As an example of this phenomenon, let us consider another type of mathematical object, namely, the set of all functions $f: \mathbf{R} \to \mathbf{R}$, which you have dealt with in calculus. We will denote this set by $F(\mathbf{R})$. $F(\mathbf{R})$ contains elements such as the functions defined by $f(x) = x^3 + 4x - 7$, $g(x) = \sin(4e^x)$, $h(x) = (|x - \cos(x)|)^{1/2}$, and so on.

We recall that if $f, g \in F(\mathbf{R})$, we can produce a new function denoted $f + g \in F(\mathbf{R})$ by defining $(f + g)(x) = f(x) + g(x)$ for all $x \in \mathbf{R}$. For instance, if $f(x) = e^x$ and $g(x) = 2e^x + x^2$, then $(f + g)(x) = 3e^x + x^2$. If $c \in \mathbf{R}$, we may multiply any $f \in F(\mathbf{R})$ by c to produce a new function denoted $cf \in F(\mathbf{R})$, and defined by $(cf)(x) = cf(x)$ for all $x \in \mathbf{R}$. Hence there are sum and scalar multiplication operations that are defined on the elements of $F(\mathbf{R})$ as well.

The analogy between algebraic operations on the vectors in $\mathbf{R}^2$ and the functions in $F(\mathbf{R})$ that is becoming apparent actually goes even deeper. We see that the following properties hold:

1. The sum operation on functions is also *associative:* For all $f, g, h \in F(\mathbf{R})$, $(f + g) + h = f + (g + h)$ as functions. This is so because for all $x \in \mathbf{R}$ we have $((f + g) + h)(x) = (f(x) + g(x)) + h(x) = f(x) + (g(x) + h(x)) = (f + (g + h))(x)$ by the associativity of addition in $\mathbf{R}$.

2. The sum operation on functions is also *commutative:* For all $f, g \in F(\mathbf{R})$, $f + g = g + f$ as functions. This follows since $f(x) + g(x) = g(x) + f(x)$ for all $x \in \mathbf{R}$. (Addition of real numbers is commutative.)

3. There is an *additive identity element* in $F(\mathbf{R})$—the constant function $z(x) = 0$. For all functions f, $f + z = f$ since $f(x) + z(x) = f(x) + 0 = f(x)$ for all $x \in \mathbf{R}$.

4. For each function f there is an *additive inverse* $-f \in F(\mathbf{R})$ (defined by $(-f)(x) = -f(x)$) with the property that $f + (-f) = z$ (the zero function from (3)).

5. For all functions f and g and all $c \in \mathbf{R}$, $c(f + g) = cf + cg$. This follows since $c(f + g)(x) = c(f(x) + g(x)) = cf(x) + cg(x) = (cf)(x) + cg(x) = (cf + cg)(x)$ by the distributivity of multiplication over addition in $\mathbf{R}$.

6. For all functions f and all $c, d \in \mathbf{R}$, $(c + d)f = cf + df$. This also follows from ordinary distributivity.

7. For all functions f and all c, $d \in \mathbf{R}$, $(cd)f = c(df)$. That is, the scalar multiple of f by the product cd is the same function as is obtained by multiplying f by d, then multiplying the result by c.

8. For all functions f, $1f = f$.

Thus, with regard to the sum and scalar multiplication operations we have defined, the elements of our two sets $\mathbf{R}^2$ and $F(\mathbf{R})$ actually behave in exactly the same way. Because of the importance and usefulness of these vectors and functions, the fact that they are so similar from an algebraic point of view is a very fortunate occurrence. Both $\mathbf{R}^2$ and $F(\mathbf{R})$ are examples of a more general mathematical structure called a vector space, which is defined as follows.

(1.1.1) Definition. A (real) *vector space* is a set V (whose elements are called *vectors* by analogy with the first example we considered) together with

a) an operation called *vector addition*, which for each pair of vectors $\mathbf{x}$, $\mathbf{y} \in V$ produces another vector in V denoted $\mathbf{x} + \mathbf{y}$, and

b) an operation called *multiplication by a scalar* (a real number), which for each vector $\mathbf{x} \in V$, and each scalar $c \in \mathbf{R}$ produces another vector in V denoted $c\mathbf{x}$.

Furthermore, the two operations must satisfy the following *axioms:*

1. For all vectors $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{z} \in V$, $(\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z})$.
2. For all vectors $\mathbf{x}$ and $\mathbf{y} \in V$, $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$.
3. There exists a vector $\mathbf{0} \in V$ with the property that $\mathbf{x} + \mathbf{0} = \mathbf{x}$ for all vectors $\mathbf{x} \in V$.
4. For each vector $\mathbf{x} \in V$, there exists a vector denoted $-x$ with the property that $\mathbf{x} + -\mathbf{x} = \mathbf{0}$.
5. For all vectors $\mathbf{x}$ and $\mathbf{y} \in V$ and all scalars $c \in \mathbf{R}$, $c(\mathbf{x} + \mathbf{y}) = c\mathbf{x} + c\mathbf{y}$.
6. For all vectors $\mathbf{x} \in V$, and all scalars c and $d \in \mathbf{R}$, $(c + d)\mathbf{x} = c\mathbf{x} + d\mathbf{x}$.
7. For all vectors $\mathbf{x} \in V$, and all scalars c and $d \in \mathbf{R}$, $(cd)\mathbf{x} = c(d\mathbf{x})$.
8. For all vectors $\mathbf{x} \in V$, $1\mathbf{x} = \mathbf{x}$.

The reason we introduce abstract definitions such as this one is that they allow us to focus on the *similarities* between different mathematical objects and treat objects with the same properties in a unified way. We will now consider several other examples of vector spaces.

(1.1.2) Example. As a first example, let us consider the set $V = \mathbf{R}^n = \{(x_1, \ldots, x_n) | x_i \in \mathbf{R} \text{ for all } i\}$. We define operations called vector addition and multiplication by a scalar by rules that are similar to the ones we saw before in $\mathbf{R}^2$.

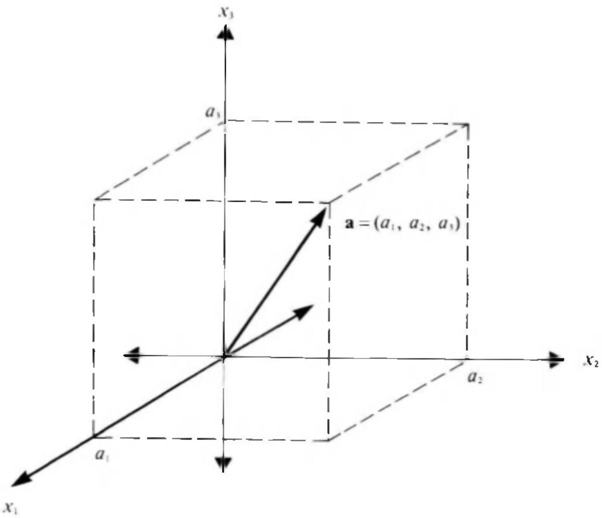

Figure 1.7

Namely, if $\mathbf{x} = (x_1, \ldots ,x_n)$ and $\mathbf{y} = (y_1, \ldots ,y_n)$, then we define $\mathbf{x} + \mathbf{y} = (x_1 + y_1, \ldots ,x_n + y_n)$ in $\mathbf{R}^n$. Next, if $\mathbf{x} = (x_1, \ldots ,x_n)$ and $c \in \mathbf{R}$, we define $c\mathbf{x} = (cx_1, \ldots ,cx_n)$. With these operations V is a vector space, since all eight axioms in the definition hold. (This may be seen by arguments that are exactly analogous to the ones given before for $\mathbf{R}^2$.) Thus we have a whole infinite collection of new vector spaces $\mathbf{R}^n$ for different integers $n \geq 1$.

When $n = 1$ we obtain $\mathbf{R}$ itself, which we may picture geometrically as a line. As we have seen, vectors in $\mathbf{R}^2$ correspond to points in the plane. Similarly, vectors in $\mathbf{R}^3$ correspond to points in a three-dimensional space such as the physical space we live in, described by Cartesian coordinates. See Figure 1.7. The vector spaces $\mathbf{R}^4$, $\mathbf{R}^5$, and so on may be thought of geometrically in a similar fashion. Due to the fact that our sense organs are adapted to an apparently three-dimensional world, it is certainly true that they are harder to visualize. Nevertheless, the reader encountering these spaces for the first time should *not* make the mistake of viewing them only as meaningless generalizations of the mathematics underlying the two- and three-dimensional physical spaces in which our geometric intuition is more at home. They are useful because frequently more than three coordinates are needed to specify a configuration or situation arising in applications of mathematics. For instance, to describe the position *and* velocity of a moving object in ordinary physical space at a given time, we actually need *six* coordinates in all (three for the position and three for the velocity, which is also a vector). Thus our "position-velocity space" may be seen as $\mathbf{R}^6$. To describe the state of the U.S. economy at a given time in a realistic economic model, hundreds or thousands of different variables or coordinates might be specified. In each of these cases, the space of all possible configurations or situations may be thought of as one of our spaces $\mathbf{R}^n$ (or a subset of one of these spaces).

(1.1.3) Example. A different kind of example of vector spaces comes from considering certain subsets of our vector space of functions. In particular, let n be a fixed nonnegative integer, and let $P_n(\mathbf{R}) = \{p: \mathbf{R} \to \mathbf{R} | p(x) = a_nx^n + a_{n-1}x^{n-1} + \cdots + a_0$, where the $a_i \in \mathbf{R}\}$, the set of all polynomial functions of a degree no larger than n. We can define the sum of two polynomial functions as we did before in the larger space $F(\mathbf{R})$. If $p(x) = a_nx^n + \cdots + a_0$ and $q(x) = b_nx^n + \cdots + b_0$, then we define

$$(p + q)(x) = p(x) + q(x) = (a_n + b_n)x^n + \cdots + (a_0 + b_0) \in P_n(\mathbf{R})$$

Similarly, if $c \in \mathbf{R}$, we define:

$$(cp)(x) = cp(x) = ca_nx^n + \cdots + ca_0 \in P_n(\mathbf{R})$$

To show that $P_n(\mathbf{R})$ is a vector space, we may verify the eight axioms given in Definition (1.1.1). These are all straightforward, and may be verified by computations that are similar to those we did before in $F(\mathbf{R})$.

(1.1.4) Examples

a) Since specifying a vector space means giving both a set of vectors and the two operations, one question that may occur to you is, given a set V, could there be more than one way to define a vector sum and a scalar multiplication to make V a vector space? For example, in $V = \mathbf{R}^2$, if we defined

$$(x_1, x_2) +' (y_1, y_2) = (x_1y_1, x_2y_2)$$

and

$$c(x_1, x_2) = (cx_1, cx_2)$$

would these operations give us another vector space?

If we look at the first three axioms, we see that they are satisfied for the new sum operation $+'$. (Do not be fooled, axiom 3 *is* satisfied if we take the identity element for $+'$ to be the vector $(1, 1)$. You should also note that this is the only vector that works here.) However, axiom 4 fails to hold here since there is no inverse for the vector $(0, 0)$ under the operation $+'$. (Why not?) Several of the other axioms fail to hold as well (which ones?). Hence we do *not* obtain a vector space in this example.

b) Similarly, we can ask if other "addition" and "scalar multiplication" operations might be defined in $F(\mathbf{R})$. Consider the addition operation $f +' g$ defined by $(f +' g)(x) = f(x) + 3g(x)$ and the usual scalar multiplication. We ask, is $F(\mathbf{R})$ a vector space with these operations? The answer is no, as the reader will see by checking the associative law for addition (axiom 1), for instance. (Other axioms fail as well. Which ones?)

c) As we know, $V = \mathbf{R}$ is a vector space. We can also make the subset $V' = \mathbf{R}^+ = \{r \in \mathbf{R} \mid r > 0\}$ into a vector space by defining our vector sum and scalar multiplication operations to be $x +' y = x \cdot y$, and $c \cdot' x = x^c$ for all x, y

$\in V'$ and all $c \in \mathbf{R}$. Since multiplication of real numbers (our *addition* operation in V') is associative and commutative, the first two axioms in Definition (1.1.1) are satisfied. The element $1 \in V'$ is an identity element for the operation $+'$, since $x +' 1 = x \cdot 1 = x$ for all $x \in V'$. Furthermore, each $x \in V'$ has an inverse $1/x \in V'$ under the operation $+'$, since $x +' (1/x) = x \cdot (1/x) = 1$.

The remaining axioms require some checking using the properties of exponents. To verify axiom 5, we compute

$$c \cdot' (x +' y) = (x \cdot y)^c = x^c \cdot y^c = x^c +' y^c = (c \cdot' x) +' (c \cdot' y)$$

Therefore axiom 5 is satisfied. The other distributive property of axiom 6 may be checked similarly. We have

$$(c+d) \cdot' x = x^{(c+d)} = x^c \cdot x^d = (c \cdot' x) +' (d \cdot' x)$$

Axiom 7 also holds since

$$(cd) \cdot' x = x^{cd} = (x^d)^c = c \cdot' (d \cdot' x)$$

Finally, $1 \cdot' x = x^1 = x$ for all $x \in V'$, so axiom 8 is satisfied as well.

We conclude this section by stating and proving some further properties of vector spaces to illustrate another benefit of introducing general definitions like Definition (1.1.1)—any proof that uses only the properties of *all* vector spaces expressed by the eight axioms is valid for all vector spaces. We do not have to reprove our results in each new example we encounter.

(1.1.5) Remark. In $\mathbf{R}^n$ there is clearly only one additive identity—the zero vector $(0, \ldots, 0) \in \mathbf{R}^n$. Moreover, each vector has only one additive inverse. We may ask whether these patterns are true in a general vector space. The fact that they are makes a general vector space somewhat simpler than it might appear at first.

(1.1.6) Proposition. Let V be a vector space. Then

a) The zero vector $\mathbf{0}$ is unique.

b) For all $\mathbf{x} \in V$, $0\mathbf{x} = \mathbf{0}$.

c) For each $\mathbf{x} \in \mathbf{V}$, the additive inverse $-x$ is unique.

d) For all $\mathbf{x} \in V$, and all $c \in \mathbf{R}$, $(-c)\mathbf{x} = -(c\mathbf{x})$.

Proof:

a) To prove that something is unique, a common technique is to assume we have two examples of the object in question, then show that those two examples must in fact be equal. So, suppose we had two vectors, $\mathbf{0}$ and $\mathbf{0}'$, both of which satisfy axiom 3 in the definition. Then, $\mathbf{0} + \mathbf{0}' = \mathbf{0}$, since $\mathbf{0}'$ is an additive identity. On the other hand, $\mathbf{0} + \mathbf{0}' = \mathbf{0}' + \mathbf{0} = \mathbf{0}'$, since addition is commutative and $\mathbf{0}$ is an additive identity. Hence $\mathbf{0} = \mathbf{0}'$, or, in other words, there is only one additive identity in V.

b) We have $0\mathbf{x} = (0 + 0)\mathbf{x} = 0\mathbf{x} + 0\mathbf{x}$, by axiom 6. Hence if we add the inverse of $0\mathbf{x}$ to both sides, we obtain $\mathbf{0} = 0\mathbf{x}$, as claimed.

c) We use the same idea as in the proof of part a. Given $\mathbf{x} \in V$, if $-\mathbf{x}$ and $(-\mathbf{x})'$ are two additive inverses of $\mathbf{x}$, then on one hand we have $\mathbf{x} + -\mathbf{x} + (-\mathbf{x})' = (\mathbf{x} + -\mathbf{x}) + (-\mathbf{x})' = \mathbf{0} + (-\mathbf{x})' = (-\mathbf{x})'$, by axioms 1, 4, and 3. On the other hand, if we use axiom 2 first before associating, we have $\mathbf{x} + -\mathbf{x} + (-\mathbf{x})' = \mathbf{x} + (-\mathbf{x})' + -\mathbf{x} = (\mathbf{x} + (-\mathbf{x})') + -\mathbf{x} = \mathbf{0} + -\mathbf{x} = -\mathbf{x}$. Hence $-\mathbf{x} = (-\mathbf{x})'$, and the additive inverse of $\mathbf{x}$ is unique.

d) We have $c\mathbf{x} + (-c)\mathbf{x} = (c + -c)\mathbf{x} = 0\mathbf{x} = \mathbf{0}$ by axiom 6 and part b. Hence $(-c)\mathbf{x}$ also serves as an additive inverse for the vector $c\mathbf{x}$. By part c, therefore, we must have $(-c)\mathbf{x} = -(c\mathbf{x})$. ■

EXERCISES

1. Let $\mathbf{x} = (1, 3, 2)$, $\mathbf{y} = (-2, 3, 4)$, $\mathbf{z} = (-3, 0, 3)$ in $\mathbf{R}^3$.
   a) Compute $3\mathbf{x}$.
   b) Compute $4\mathbf{x} - \mathbf{y}$.
   c) Compute $-\mathbf{x} + \mathbf{y} + 3\mathbf{z}$.

2. Let $f = 3e^{3x}$, $g = 4e^{3x} + e^x$, $h = 2e^x - e^{3x}$ in $F(\mathbf{R})$.
   a) Compute $5f$.
   b) Compute $2f + 3g$.
   c) Compute $-2f - g + 4h$.

3. Show that $\mathbf{R}^n$ with the vector sum and scalar multiplication operations given in Example (1.1.2) is a vector space.

4. Complete the proof that $P_n(\mathbf{R})$, with the operations given in Example (1.1.3), is a vector space.

5. Let $V = \{p: \mathbf{R} \to \mathbf{R} \mid p(x) = a_nx^n + \cdots + a_0$, where the $a_i \in \mathbf{R}$, and $a_n \neq 0\}$ (the set of polynomial functions of degree exactly n). Is V a vector space, using the operations given in Example (1.1.3)? Why or why not?

6. In each of the following parts, decide if the set $\mathbf{R}^2$, with the given operations, is a vector space. If this is not the case, say which of the axioms fail to hold.
   a) vector sum $(x_1, x_2) +' (y_1, y_2) = (x_1 + 2y_1, 3x_2 - y_2)$, and the usual scalar multiplication $c(x_1, x_2) = (cx_1, cx_2)$
   b) usual vector sum $(x_1, x_2) + (y_1, y_2) = (x_1 + y_1, x_2 + y_2)$, and scalar multiplication $c(x_1, x_2) = \begin{cases} (cx_1, (1/c)x_2) \text{ if } c \neq 0 \\ (0, 0) \text{ if } c = 0 \end{cases}$
   c) vector sum $(x_1, x_2) +' (y_1, y_2) = (0, x_1 + y_2)$, and the usual scalar multiplication

7. In each of the following parts, decide if the set $F(\mathbf{R})$, with the given operations, is a vector space. If this is not the case, say which of the axioms fail to hold.
   a) Sum operation defined by $f +' g = fg$, scalar multiplication given by $c \cdot f = c + f$, that is, the constant function c plus f

b) Sum defined by $f +' g = f - g$, scalar multiplication given by $(c \cdot f)(x) = f(cx)$
c) Sum defined by $f +' g = f \circ g$ (composition of functions), usual scalar multiplication

8. Show that in any vector space V
a) If $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$, then $\mathbf{x} + \mathbf{y} = \mathbf{x} + \mathbf{z}$ implies $\mathbf{y} = \mathbf{z}$.
b) If $\mathbf{x}, \mathbf{y} \in V$ and $a, b \in \mathbf{R}$, then $(a + b)(\mathbf{x} + \mathbf{y}) = a\mathbf{x} + b\mathbf{x} + a\mathbf{y} + b\mathbf{y}$.

9. a) What vector space might be used to describe (simultaneously) the position, velocity, and acceleration of an object moving along a path in the plane $\mathbf{R}^2$?
b) Same question for an object moving in three-dimensional space.

10. Let $[a, b]$ be the closed interval $\{x \in \mathbf{R} \mid a \leq x \leq b\} \subset \mathbf{R}$. Let $F([a, b])$ be the set of all functions $f: [a, b] \to \mathbf{R}$. Show that $F([a,b])$ is a vector space if we define the sum and scalar multiplication operations as in $F(\mathbf{R})$.

11. Let $V = \{a_1x^2 + a_2xy + a_3y^2 + a_4x + a_5y + a_6 \mid a_i \in \mathbf{R}\}$, the set of all polynomials in two variables x and y of total degree no larger than 2. Define sum and scalar multiplication operations in V as in the vector space $P_n(\mathbf{R})$; that is, the sum operation is ordinary addition of polynomials, and multiplication by a scalar multiplies each coefficient by that scalar. Show that V, with these operations, is a vector space.

12. Let $V = (\mathbf{R}^+)^n = \{(x_1, \ldots, x_n) \mid x_i \in \mathbf{R}^+ \text{ for each } i\}$. See Example (1.1.4c). In V define a vector sum operation $+'$ by $(x_1, \ldots, x_n) +' (y_1, \ldots, y_n) = (x_1y_1, \ldots, x_ny_n)$, and a scalar multiplication operation $\cdot'$ by $c \cdot' (x_1, \ldots, x_n) = (x_1^c, \ldots, x_n^c)$. Show that with these two operations V is a vector space.

§1.2. SUBSPACES

In Section 1.1 we saw several different types of examples of vector spaces, but those examples are far from a complete list of the vector spaces that arise in different areas of mathematics. In this section we begin by indicating some other important examples. Another type of example, vector spaces of matrices, is introduced in the exercises following this section.

Our first example deals with a special kind of subset of the vector space $\mathbf{R}^3$, which we introduced in Example (1.1.2).

(1.2.1) Example. Let $V = \{(x_1, x_2, x_3) \in \mathbf{R}^3 \mid 5x_1 - 2x_2 + x_3 = 0\}$. Geometrically, the set V is a plane passing through the origin in $\mathbf{R}^3$. See Figure 1.8. If we endow V with the operations of vector addition and scalar multiplication from the space $\mathbf{R}^3$, then V is also a vector space. Note first that if $\mathbf{x} = (x_1, x_2, x_3)$ and $\mathbf{y} = (y_1, y_2, y_3) \in V$, then the sum $\mathbf{x} + \mathbf{y} = (x_1 + y_1, x_2 + y_2, x_3 + y_3) \in V$

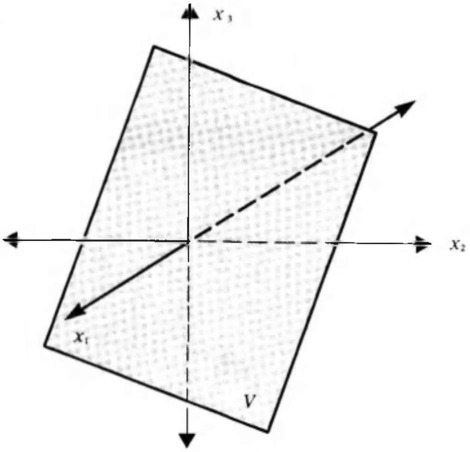

Figure 1.8

as well. This follows because the components of the sum also satisfy the defining equation of V:

$$\begin{aligned} 5(x_1 + y_1) - 2(x_2 + y_2) + (x_3 + y_3) &= (5x_1 - 2x_2 + x_3) + (5y_1 - 2y_2 + y_3) \\ &= 0 + 0 \\ &= 0 \end{aligned}$$

Similarly, for any $\mathbf{x} \in V$, and any $c \in \mathbf{R}$, $c\mathbf{x} \in V$, since again the components of the vector $c\mathbf{x}$ satisfy the defining equation of V:

$$5(cx_1) - 2(cx_2) + (cx_3) = c(5x_1 - 2x_2 + x_3) = 0$$

The zero vector (0, 0, 0) also satisfies the defining equation of V.

Now, since V is a subset of the vector space $\mathbf{R}^3$ and the vector sum and scalar multiplication are defined in the same way in both V and in $\mathbf{R}^3$, then it follows that the axioms for vector spaces are satisfied by V as well. Since they hold for all vectors in $\mathbf{R}^3$, they also hold for the vectors in the subset V.

(1.2.2) Example. The previous example may be generalized immediately. In $\mathbf{R}^n$, consider any set defined in the following way. Let $V = \{(x_1, \ldots, x_n) \in \mathbf{R}^n \mid a_1x_1 + \cdots + a_nx_n = 0$, where $a_i \in \mathbf{R}$ for all $i\}$. Then V is a vector space, if we define the vector sum and scalar multiplication to be the same as the operations in the whole space $\mathbf{R}^n$. V is sometimes called a hyperplane in $\mathbf{R}^n$.

(1.2.3) Example. Many of the sets of functions you have encountered in calculus give further examples of vector spaces. For example, consider the set $V = \{f\colon\ \mathbf{R} \to \mathbf{R}$

$\mid f$ is continuous}. We usually denote this set of functions by $C(\mathbf{R})$. Recall that f is continuous if and only if for all $a \in \mathbf{R}$, we have $\lim_{x \to a} f(x) = f(a)$. We define vector sum and scalar multiplication in $C(\mathbf{R})$ as usual for functions. If $f, g \in C(\mathbf{R})$ and $c \in \mathbf{R}$, then $f + g$ is the function with $(f + g)(x) = f(x) + g(x)$ for all $x \in \mathbf{R}$, and cf is the function $(cf)(x) = cf(x)$ for all $x \in \mathbf{R}$. We claim that the set $C(\mathbf{R})$ with these two operations is a vector space.

To see why this is true, we need to make some preliminary observations. In our previous examples of vector spaces, such as $\mathbf{R}^n$, $F(\mathbf{R})$ and $P_n(\mathbf{R})$, it was more or less clear from the definitions that applying the appropriate vector sum and scalar multiplication operations to vectors in the set in question gave us elements of the same set, a requirement that is part of the definition of a vector space. The fact that this is also true in this example is less trivial. What is involved here is a pair of important properties of continuous functions, which we summarize in the following result.

(1.2.4) Lemma. Let $f, g \in C(\mathbf{R})$, and let $c \in \mathbf{R}$. Then

a) $f + g \in C(\mathbf{R})$, and

b) $cf \in C(\mathbf{R})$.

Proof:

a) By the limit sum rule from calculus, for all $a \in \mathbf{R}$ we have

$$\lim_{x \to a} (f + g)(x) = \lim_{x \to a} (f(x) + g(x)) = \lim_{x \to a} f(x) + \lim_{x \to a} g(x)$$

Since f and g are continuous, this last expression is equal to $f(a) + g(a) = (f + g)(a)$. Hence $f + g$ is continuous.

b) By the limit product rule, we have

$$\lim_{x \to a} (cf)(x) = \lim_{x \to a} cf(x) = (\lim_{x \to a} c) \cdot (\lim_{x \to a} f(x)) = cf(a) = (cf)(a)$$

so cf is also continuous. ■

That the eight axioms in Definition (1.1.1) hold in $C(\mathbf{R})$ may now be verified in much the same way that we verified them for the vector space $F(\mathbf{R})$ in Section 1.1. Alternately, we may notice that these verifications are actually unnecessary—since we have already established those properties for all the functions in $F(\mathbf{R})$, they must also hold for the functions in the subset $C(\mathbf{R})$. Of course, we must convince ourselves that the zero function from $F(\mathbf{R})$ also serves as an additive identity element for $C(\mathbf{R})$, but this is clear.

(1.2.5) Example. Now, let $V = \{f: \mathbf{R} \to \mathbf{R} \mid f$ differentiable everywhere, and $f'(x) \in C(\mathbf{R})\}$. The elements of V are called continuously differentiable functions and the set V is usually denoted by $C^1(\mathbf{R})$. If we define sums of functions and scalar multiples as in $F(\mathbf{R})$, then $C^1(\mathbf{R})$ is also a vector space.

To see that this is true, once again we should start by checking that sums and scalar multiples of functions in $C^1(\mathbf{R})$ are indeed in $C^1(\mathbf{R})$. This follows from properties of differentiation that you learned in calculus. Note that if $f, g \in C^1(\mathbf{R})$, the sum rule for derivatives implies that $f + g$ is also differentiable and that $(f + g)' = f' + g'$. Then, since both f' and g' are in $C(\mathbf{R})$, by part (a) of Lemma (1.2.4) we have that $f' + g' \in C(\mathbf{R})$. Hence $f + g \in C^1(\mathbf{R})$. Similarly, if $f \in C^1(\mathbf{R})$ and $c \in \mathbf{R}$, then the product rule for derivatives implies that cf is differentiable and $(cf)' = cf'$. Then, by part (b) of Lemma (1.2.4), since $f' \in C(\mathbf{R})$, we see that $cf' \in C(\mathbf{R})$. Hence $cf \in C^1(\mathbf{R})$ as well.

It is clear that the zero function is continuously differentiable so $C^1(\mathbf{R})$ does have an additive identity element. The fact that the remaining axioms for vector spaces hold in $C^1(\mathbf{R})$ is now actually a direct consequence of our previous verifications in $C(\mathbf{R})$ or $F(\mathbf{R})$. Since those properties hold for all functions in $F(\mathbf{R})$, they also hold for the functions in the subset $C^1(\mathbf{R})$.

In all these examples, you should note that we were dealing with subsets of vector spaces that were vector spaces in their own right. In general, we will use the following definition.

(1.2.6) Definition. Let V be a vector space and let $W \subseteq V$ be a subset. Then W is a (vector) *subspace* of V if W is a vector space itself under the operations of vector sum and scalar multiplication from V.

(1.2.7) Examples

a) From Example (1.2.2), for each particular set of coefficients $a_i \in \mathbf{R}$, the vector space $W = \{(x_1, \ldots, x_n) \in \mathbf{R}^n \mid a_1x_1 + \cdots + a_nx_n = 0\}$ is a subspace of $\mathbf{R}^n$.

b) $C(\mathbf{R})$, $P_n(\mathbf{R})$, $C^1(\mathbf{R})$ are all subspaces of $F(\mathbf{R})$.

c) $C^1(\mathbf{R})$ is a subspace of $C(\mathbf{R})$, since differentiability implies continuity.

d) For each n, the vector space $P_n(\mathbf{R})$ is a subspace of $C^1(\mathbf{R})$.

e) The vector space structure we defined on the subset $\mathbf{R}^+ \subset \mathbf{R}$ in Example (1.1.4c) does *not* make $\mathbf{R}^+$ a subspace of $\mathbf{R}$. The reason is that the vector sum and scalar multiplication operations in $\mathbf{R}^+$ are different from those in $\mathbf{R}$, while in a subspace, we must use the same operations as in the larger vector space.

f) On a more general note, in every vector space, the subsets V and $\{\mathbf{0}\}$ are subspaces. This is clear for V itself. It is also easy to verify that the set containing the one vector $\mathbf{0}$ is a subspace of V.

Although it is always possible to determine if a subset W of a vector space is a subspace directly [by checking the axioms from Definition (1.1.1)], it would be desirable to have a more economical way of doing this. The observations we have made in the previous examples actually indicate a general criterion that can be used to tell when a subset of a vector space is a subspace. Note that if W is to be a vector space in its own right, then the following properties must be true:

1. For all $\mathbf{x}, \mathbf{y} \in W$, we must have $\mathbf{x} + \mathbf{y} \in W$. If W has this property, we say that the set W is *closed under addition*.

2. For all $\mathbf{x} \in W$ and all $c \in \mathbf{R}$, we must have $c\mathbf{x} \in W$. If W has this property, we say that W is *closed under scalar multiplication*.

3. The zero vector of V must be contained in W.

Note that since the vector sum operation is the same in W as in V, the additive identity element in W must be the same as it is in V, by Proposition (1.1.6a). In fact, it is possible to *condense* these three conditions into a single condition that is easily checked and that characterizes the subspaces of V completely.

(1.2.8) Theorem. Let V be a vector space, and let W be a nonempty subset of V. Then W is a subspace of V if and only if for all $\mathbf{x}, \mathbf{y} \in W$, and all $c \in \mathbf{R}$, we have $c\mathbf{x} + \mathbf{y} \in W$.

(1.2.9) Remarks

a) By Definition (1.1.1) a vector space must contain at least an additive identity element, hence the requirement that W be nonempty is certainly necessary.

b) Before we begin the proof itself, notice that any statement of the form "p if and only if q" is equivalent to the statement "if p then q and if q then p." (The "only if" part is the statement "if p then q.") Thus, to prove an if and only if statement, we must prove both the *direct implication* "if p then q," and the *reverse implication* or *converse* "if q then p." To signal these two sections in the proof of an if and only if statement, we use the symbol $\rightarrow$ to indicate proof of the direct implication, and the symbol $\leftarrow$ to indicate the proof of the reverse implication. (The reader may wish to consult Appendix 1, section b for further comments on the logic involved here.)

Proof: $\rightarrow$: If W is a subspace of V, then for all $\mathbf{x} \in W$, and all $c \in \mathbf{R}$, we have $c\mathbf{x} \in W$, and hence for all $\mathbf{y} \in W$, $c\mathbf{x} + \mathbf{y} \in W$ as well, because by the definition a subspace W of V must be closed under vector sums and scalar multiples.

$\leftarrow$: Let W be any subset of V satisfying the condition of the theorem. First, note that since $c\mathbf{x} + \mathbf{y} \in W$ for all choices of $\mathbf{x}, \mathbf{y} \in W$ and $c \in \mathbf{R}$, we may specialize to the case $c = 1$. Then we see that $1\mathbf{x} + \mathbf{y} = \mathbf{x} + \mathbf{y} \in W$, so that W is closed under sums. Next, let $\mathbf{x} = \mathbf{y}$ be any vector in W and $c = -1$. Then $-1\mathbf{x} + \mathbf{x} = (-1 + 1)\mathbf{x} = 0\mathbf{x} = \mathbf{0} \in W$. Now let $\mathbf{x}$ be any vector in W and let $\mathbf{y} = 0$. Then $c\mathbf{x} + \mathbf{0} = c\mathbf{x} \in W$, so W is closed under scalar multiplication. To see that these observations imply that W is a vector space, note that the axioms 1, 2, and 5 through 8 in Definition (1.1.1) are satisfied automatically for vectors in W, since they hold for all vectors in V. Axiom 3 is satisfied, since as we have seen $\mathbf{0} \in W$. Finally, for each $\mathbf{x} \in W$, by Proposition (1.1.6d) $(-1)\mathbf{x} = -\mathbf{x} \in W$ as well. Hence W is a vector space. ∎

To see how the condition of the theorem may be applied, we consider several examples.

(1.2.10) Example. In $V = \mathbf{R}^3$ consider the subset

$$W = \{(x_1, x_2, x_3) \mid 4x_1 + 3x_2 - 2x_3 = 0, \text{ and } x_1 - x_3 = 0\}$$

By Theorem (1.2.8), W is a subspace of $\mathbf{R}^3$ since if $\mathbf{x}, \mathbf{y} \in W$ and $c \in \mathbf{R}$, then writing $\mathbf{x} = (x_1, x_2, x_3)$ and $\mathbf{y} = (y_1, y_2, y_3)$, we have that the components of the vector $c\mathbf{x} + \mathbf{y} = (cx_1 + y_1, cx_2 + y_2, cx_3 + y_3)$ satisfy the defining equations of the set W:

$$\begin{aligned} 4(cx_1 + y_1) + 3(cx_2 + y_2) - 2(x_3 + y_3) &= c(4x_1 + 3x_2 - 2x_3) + \\ &\qquad (4y_1 + 3y_2 - 2y_3) \\ &= c0 + 0 = 0 \end{aligned}$$

Similarly,

$$\begin{aligned} (cx_1 + y_1) - (cx_3 + y_3) &= c(x_1 - x_3) + (y_1 - y_3) \\ &= c0 + 0 = 0 \end{aligned}$$

Hence $c\mathbf{x} + \mathbf{y} \in W$. W is nonempty since the zero vector $(0, 0, 0)$ satisfies both equations.

(1.2.11) Examples

a) In $V = C(\mathbf{R})$, consider the set $W = \{f \in C(\mathbf{R}) \mid f(2) = 0\}$. First, W is nonempty, since it contains functions such as $f(x) = x - 2$. W is a subspace of V, since if $f, g \in W$, and $c \in \mathbf{R}$, then we have $(cf + g)(2) = cf(2) + g(2) = c0 + 0 = 0$. Hence $cf + g \in W$ as well.

b) Let $C^2(\mathbf{R})$ denote the set of functions $f \in F(\mathbf{R})$ such that f is twice differentiable and $f'' \in C(\mathbf{R})$. We will show that $C^2(\mathbf{R})$ is a subspace of $C(\mathbf{R})$. First, since every differentiable function on $\mathbf{R}$ is continuous everywhere, we have that $C^2(\mathbf{R}) \subset C(\mathbf{R})$. Furthermore, $C^2(\mathbf{R})$ is nonempty since it certainly contains all polynomial functions. We will show that the criterion given in Theorem (1.2.8) is satisfied. Let f, g be any functions in $C^2(\mathbf{R})$ and consider the function $cf + g$. What we must show is that $cf + g$ is also twice differentiable with continuous second derivative.

First, since f and g can be differentiated twice, $cf + g$ is also twice differentiable, and the sum and scalar product rules for derivatives show that $(cf + g)'' = (cf' + g')' = cf'' + g''$. Second, since f'' and g'' are continuous functions, by Lemma (1.2.4), $cf'' + g''$ is also continuous. Therefore $cf + g \in C^2(\mathbf{R})$, so $C^2(\mathbf{R})$ is a subspace of $C(\mathbf{R})$.

Since you may not have seen examples of this kind before, we include an example of a function $f \in C^1(\mathbf{R})$ which is *not* in $C^2(\mathbf{R})$. This shows that the vector space $C^2(\mathbf{R})$ is contained in, but is not equal to $C^1(\mathbf{R})$. Let

$$f(x) = \begin{cases} x^2 & \text{if } x > 0 \\ -x^2 & \text{if } x \leq 0 \end{cases}.$$

Then f is differentiable everywhere, and

$$f'(x) = \begin{cases} 2x \text{ if } x > 0 \\ 0 \text{ if } x = 0 \\ -2x \text{ if } x < 0 \end{cases}.$$

(You should check that in fact $f'(0) = \lim_{h \to 0} [f(0+h) - f(0)]/h = 0$, as we claim.) Furthermore, f' is a continuous function, which shows that $f \in C^1(\mathbf{R})$. However, f' is not differentiable at $x = 0$, since the graph of f' has a corner there—indeed, the formulas for f' given above show that f' is the same function as $g(x) = 2|x|$. Hence $f \notin C^2(\mathbf{R})$.

Theorem (1.2.8) may also be used to show that subsets of vector spaces are not subspaces.

(1.2.12) Example. In $V = \mathbf{R}^2$, consider $W = \{(x_1, x_2) \mid x_1^3 - x_2^2 = 0\}$. This W is not a subspace of $\mathbf{R}^2$, since, for instance, we have $(1, 1)$ and $(4, 8) \in W$, but the sum $(1, 1) + (4, 8) = (5, 9) \notin W$. The components of the sum do not satisfy the defining equation of W: $5^3 - 9^2 = 44 \neq 0$.

Let us return now to the subspace W of $\mathbf{R}^3$ given in Example (1.2.10). Note that if we define

$$W_1 = \{(x_1, x_2, x_3) \in \mathbf{R}^3 \mid 4x_1 + 3x_2 - 2x_3 = 0\}$$

and

$$W_2 = \{(x_1, x_2, x_3) \in \mathbf{R}^3 \mid x_1 - x_3 = 0\}$$

then W is the set of vectors in both W_1 and W_2. In other words, we have an equality of sets $W = W_1 \cap W_2$. In this example we see that the *intersection* of these two subspaces of $\mathbf{R}^3$ is also a subspace of $\mathbf{R}^3$. This is a property of intersections of subspaces, which is true in general.

(1.2.13) Theorem. Let V be a vector space. Then the intersection of any collection of subspaces of V is a subspace of V.

Proof: Consider any collection of subspaces of V. Note first that the intersection of the subspaces is nonempty, since it contains at least the zero vector from V. Now, let $\mathbf{x}$, $\mathbf{y}$ be any two vectors in the intersection of all the subspaces in the collection (i.e., $\mathbf{x}, \mathbf{y} \in W$ for all W in the collection). Since each W in the collection is a subspace of V, $c\mathbf{x} + \mathbf{y} \in W$. Since this is true for all the W in the collection, $c\mathbf{x} + \mathbf{y}$ is in the intersection of all the subspaces in the collection. Hence the intersection is a subspace of V by Theorem (1.2.8). ■

One important application of this theorem deals with general subspaces of the space $\mathbf{R}^n$ of the form seen in Example (1.2.10). Namely, we will show that the set of all solutions of any *simultaneous system* of equations of the form

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= 0 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= 0 \end{aligned}$$

is a subspace of $\mathbf{R}^n$. (Here the notation a_{ij} means the coefficient of x_j in the ith equation. For example, a_{23} is the coefficient of x_3 in the second equation in the system.)

(1.2.14) Corollary. Let a_{ij} ($1 \leqslant i \leqslant m$, $1 \leqslant j \leqslant n$) be any real numbers and let $W = \{(x_1, \ldots, x_n) \in \mathbf{R}^n \mid a_{i1}x_1 + \cdots + a_{in}x_n = 0 \text{ for all } i,\ 1 \leqslant i \leqslant m\}$. Then W is a subspace of $\mathbf{R}^n$.

Proof: For each i, $1 \leqslant i \leqslant m$, let $W_i = \{(x_1, \ldots, x_n) \mid a_{i1}x_1 + \cdots + a_{in}x_n = 0\}$. Then since W is precisely the set of solutions of the simultaneous system formed from the defining equations of all the W_i, we have $W = W_1 \cap W_2 \cap \cdots \cap W_m$. Each W_i is a subspace of $\mathbf{R}^n$ [see Example (1.2.2)], so by Theorem (1.2.13) W is also a subspace of $\mathbf{R}^n$. ■

Historically, finding methods for solving systems of equations of this type was the major impetus for the development of linear algebra. In addition, many of the applications of linear algebra come down to solving systems of linear equations (and describing the set of solutions in some way). We will return to these matters in Section 1.5 of this chapter.

EXERCISES

1. Show that $\{\mathbf{0}\} \subset V$ is a subspace of each vector space V.

2. a) Let $V_1 = \{f: \mathbf{R} \to \mathbf{R} \mid f(x) = f(-x) \text{ for all } x \in \mathbf{R}\}$. ($V_1$ is called the set of *even* functions.) Show that $\cos(x)$ and x^2 define functions in V_1.
   b) Show that V_1 is a subspace of $F(\mathbf{R})$ using the same operations given in Example (1.2.1).
   c) Let $V_2 = \{f: \mathbf{R} \to \mathbf{R} \mid f(-x) = -f(x) \text{ for all } x \in \mathbf{R}\}$. ($V_2$ is called the set of *odd* functions.) Give three examples of functions in V_2.
   d) Show that V_2 is also a subspace of $F(\mathbf{R})$.

3. For each of the following subsets W of a vector space V, determine if W is a subspace of V. Say why or why not in each case:
   a) $V = \mathbf{R}^3$, and $W = \{(a_1, a_2, a_3) \mid a_1 - 3a_2 + 4a_3 = 0, \text{ and } a_1 = a_2\}$
   b) $V = \mathbf{R}^2$, and $W = \{(a_1, a_2) \mid \sin(a_1) = a_2\}$
   $V = \mathbf{R}^3$, and $W = \{(a_1, a_2, a_3) \mid (a_1 + a_2 + a_3)^2 = 0\}$

d) $V = \mathbf{R}^3$, and $W = \{(a_1, a_2, a_3) \mid a_3 \geq 0\}$
e) $V = \mathbf{R}^3$, and $W = \{(a_1, a_2, a_3) \mid a_1, a_2, a_3 \text{ all integers}\}$
f) $V = C^1(\mathbf{R})$, and $W = \{f \mid f'(x) + 4f(x) = 0 \text{ for all } x \in \mathbf{R}\}$
g) $V = C^1(\mathbf{R})$, and $W = \{f \mid \sin(x) \cdot f'(x) + f(x) = 6 \text{ for all } x \in \mathbf{R}\}$
h) $V = P_n(\mathbf{R})$, and $W = \{p \mid p(\sqrt{2}) = 0\}$
i) $V = P_n(\mathbf{R})$, and $W = \{p \mid p(1) = 1 \text{ and } p(2) = 0\}$.
j) $V = P_3(\mathbf{R})$, and $W = \{p \mid p'(x) \in P_1(\mathbf{R})\}$
k) $V = F(\mathbf{R})$, and $W = \{f \mid f \text{ is periodic with period } 2\pi\text{: } f(x + 2\pi) = f(x)$ for all $x \in \mathbf{R}\}$

4. a) If W is a subspace of a vector space V, show that for all vectors $\mathbf{x}_1, \ldots, \mathbf{x}_n \in W$, and all scalars $a_1, \ldots, a_n \in \mathbf{R}$, the vector $a_1\mathbf{x}_1 + \cdots + a_n\mathbf{x}_n \in W$.
   b) Is the converse of the statement in part a true?

5. Let W be a subspace of a vector space V, let $\mathbf{y} \in V$, and define the set $\mathbf{y} + W = \{\mathbf{x} \in V \mid \mathbf{x} = \mathbf{y} + \mathbf{w} \text{ for some } \mathbf{w} \in W\}$. Show that $\mathbf{y} + W$ is a subspace of V if and only if $\mathbf{y} \in W$.

6. If W_1 and W_2 are subspaces of a vector space V, is $W_1 \setminus W_2$ ever a subspace of V? Why or why not? (Here $W_1 \setminus W_2$ denotes the *set difference* of W_1 and W_2: $W_1 \setminus W_2 = \{\mathbf{w} \in W \mid \mathbf{w} \in W_1 \text{ but } \mathbf{w} \notin W_2\}$.)

7. a) Show that in $V = \mathbf{R}^2$, each line containing the origin is a subspace.
   b) Show that the only subspaces of $V = \mathbf{R}^2$ are the zero subspace, $\mathbf{R}^2$ itself, and the lines through the origin. (*Hint:* Show that if W is a subspace of $\mathbf{R}^2$ that contains two nonzero vectors lying along different lines through the origin, then W must be all of $\mathbf{R}^2$.)

8. Let $C([a, b])$ denote the set of continuous functions on the closed interval $[a, b] \subset \mathbf{R}$. Show that $C([a, b])$ is a subspace of the vector space $F([a, b])$ introduced in Exercise 10 of Section 1.1.

9. Let $C^\infty(\mathbf{R})$ denote the set of functions in $F(\mathbf{R})$ that have derivatives of all orders. Show that $C^\infty(\mathbf{R})$ is a subspace of $F(\mathbf{R})$.

10. Show that if V_1 is a subspace of V_2 and V_2 is a subspace of V_3, then V_1 is a subspace of V_3.

The following group of exercises introduces new examples of vector spaces that will be used extensively later in the text. Let $m, n \geq 1$ be integers. An *m by n matrix* is a rectangular array of real numbers with m (horizontal) rows and n (vertical) columns:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}.$$

Here a_{ij} represents the entry in the ith row and the jth column of the matrix. The set of all m by n matrices with real entries will be denoted by $M_{m\times n}(\mathbf{R})$. We usually use the shorthand notation $A = (a_{ij})$ to indicate the matrix whose entries are the a_{ij}.

If $A = (a_{ij})$ and $B = (b_{ij})$ are both matrices in $M_{m\times n}(\mathbf{R})$, we can define their *sum*, denoted $A + B$, to be the m by n matrix whose entries are the sums of the corresponding entries from A and B. That is, $A + B$ is the matrix whose entries are $(a_{ij} + b_{ij})$. (This sum operation is not defined if the matrices A and B have different sizes.) In addition, given a matrix $A = (a_{ij}) \in M_{m\times n}(\mathbf{R})$ and a scalar $c \in \mathbf{R}$, we define the product of c and A to be the matrix cA whose entries are obtained by multiplying each entry of A by c: $cA = (ca_{ij})$.

11. Using the definitions of the matrix sum and scalar product operations given earlier, compute:

a) $\begin{bmatrix} 2 & 4 & -2 \\ 3 & 1 & 9 \end{bmatrix} + \begin{bmatrix} -1 & 1 & 0 \\ 7 & 1 & 4 \end{bmatrix}$ in $M_{2\times 3}(\mathbf{R})$

b) $4\cdot\begin{bmatrix} 2 & -2 \\ 6 & -1 \end{bmatrix} + 3\cdot\begin{bmatrix} -2 & -9 \\ 1 & 4 \end{bmatrix} - 2\cdot\begin{bmatrix} 1 & 1 \\ 5 & -3 \end{bmatrix}$ in $M_{2\times 2}(\mathbf{R})$

12. Show that $M_{m\times n}(\mathbf{R})$ is a vector space, using the sum and scalar product operations defined earlier.

13. Show that the subset $W = \{\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \in M_{2\times 2}(\mathbf{R}) \mid 3a_{11} - 2a_{22} = 0\}$ is a subspace of $M_{2\times 2}(\mathbf{R})$.

14. Show that the subset $W = \{\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \in M_{2\times 2}(\mathbf{R}) \mid a_{12} = a_{21}\}$ (called the set of *symmetric* 2 by 2 matrices) is a subspace of $M_{2\times 2}(\mathbf{R})$.

15. Show that the subset $W = \{\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \in M_{2\times 2}(\mathbf{R}) \mid a_{11} = a_{22} = 0$ and $a_{12} = -a_{21}\}$ (called the set of *skew-symmetric* 2 by 2 matrices) is a subspace of $M_{2\times 2}(\mathbf{R})$.

16. Ignoring the way the elements are written, do you see any similarities between the vector space $M_{m\times n}(\mathbf{R})$ and other vector spaces we have studied? Try to construct a one-to-one correspondence between the m by n matrices and the elements of another vector space.

§1.3. LINEAR COMBINATIONS

If we apply the operations of vector addition and multiplication by scalars repeatedly to vectors in a vector space V, the most general expressions we can produce have the form $a_1\mathbf{x}_1 + \cdots + a_n\mathbf{x}_n$, where the $a_i \in \mathbf{R}$, and the $\mathbf{x}_i \in V$. Frequently, the vectors involved will come from some specified subset of V. To discuss this situation, we introduce the following terminology.

(1.3.1) Definitions. Let S be a subset of a vector space V.

a) A *linear combination* of vectors in S is any sum $a_1\mathbf{x}_1 + \cdots + a_n\mathbf{x}_n$, where the $a_i \in \mathbf{R}$, and the $\mathbf{x}_i \in S$.

b) If $S \neq \phi$ (the empty subset of V), the set of all linear combinations of vectors in S is called the (linear) *span* of S, and denoted Span(S). If $S = \phi$, we define Span(S) $= \{\mathbf{0}\}$.

c) If $W =$ Span(S), we say S *spans* (or *generates*) W.

We think of the span of a set S as the set of all vectors that can be "built up" from the vectors in S by forming linear combinations.

(1.3.2) Example. In $V = \mathbf{R}^3$, let $S = \{(1, 0, 0), (0, 1, 0)\}$. Then a typical linear combination of the vectors in S is a vector

$$a_1(1, 0, 0) + a_2(0, 1, 0) = (a_1, a_2, 0)$$

The span of S is the set of all such vectors (i.e., the vectors produced for all choices of $a_1, a_2 \in \mathbf{R}$). We have Span $(S) = \{(a_1, a_2, 0) \in \mathbf{R}^3 \mid a_1, a_2 \in \mathbf{R}\}$. Geometrically, Span($S$) is just the $x_1 - x_2$-plane in $\mathbf{R}^3$. See Figure 1.9. Note that in this example, we can also describe Span(S) as the set of all vectors in $\mathbf{R}^3$ whose third components are 0, that is, Span(S) $= \{(a_1, a_2, a_3) \in \mathbf{R}^3 \mid a_3 = 0\}$. Hence, by Corollary (1.2.14), Span(S) is a subspace of $\mathbf{R}^3$.

(1.3.3) Example. In $V = C(\mathbf{R})$, let $S = \{1, x, x^2, \ldots, x^n\}$. Then we have Span($S$) $= \{f \in C(\mathbf{R}) \mid f(x) = a_0 + a_1x + \cdots + a_nx^n$ for some $a_0, \ldots, a_n \in \mathbf{R}\}$. Thus Span($S$) is the subspace $P_n(\mathbf{R}) \subset C(\mathbf{R})$.

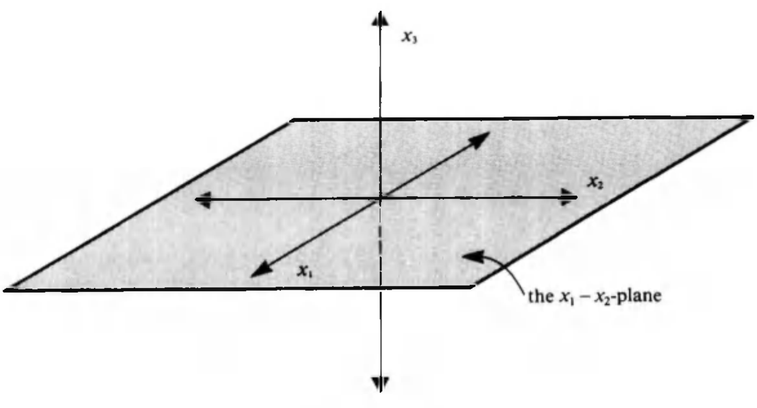

Figure 1.9.