



## Chapter 8

# Functions on Euclidean Space

In this chapter we begin the study of functions defined on a subset of the Euclidean space  $\mathbb{R}^p$  with values in the Euclidean space  $\mathbb{R}^q$ . Our first objective is to define and study continuity for such functions.

### 8.1 Continuous Functions of Several Variables

For two natural numbers  $p$  and  $q$ , we shall study functions  $F$ , defined on a subset  $D$  of  $\mathbb{R}^p$  and with values in  $\mathbb{R}^q$ . Such a function is sometimes called a *transformation* from  $D$  to  $\mathbb{R}^q$ . We will denote this situation by  $F : D \rightarrow \mathbb{R}^q$ . The definition of continuity in this context follows the familiar pattern.

**Definition 8.1.1.** Let  $D$  be a subset of  $\mathbb{R}^p$  and  $F : D \rightarrow \mathbb{R}^q$  a function. We say that  $F$  is continuous at  $a \in D$  if for each  $\epsilon > 0$  there is a  $\delta > 0$  such that

$$\|F(x) - F(a)\| \leq \epsilon \quad \text{whenever } x \in D \quad \text{and} \quad \|x - a\| < \delta.$$

If  $F$  is continuous at each point of  $D$ , then  $F$  is said to be continuous on  $D$ .

Note that this definition depends very much on the domain  $D$  of the function due to the fact that the condition on  $\|F(x) - F(a)\|$  is only required to hold for  $x \in D$ . If the domain of the function is changed, then what it means for a function to be continuous at  $a$  may change even if  $a$  is in both domains.

**Example 8.1.2.** The function  $f : \mathbb{R}^p \rightarrow \mathbb{R}$  which is 1 on  $\overline{B}_1(0)$  and 0 everywhere else is clearly not continuous at boundary points of  $\overline{B}_1(0)$ . Show that, if the domain of  $f$  is changed to  $\overline{B}_1(0)$ , then the new function is continuous on all of  $\overline{B}_1(0)$ .

**Solution:** The new function is just the identically 1 function on its domain and, hence, is continuous at each point of its domain – including points of the boundary.

**Example 8.1.3.** Consider the function  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  defined by

$$f(x, y) = \begin{cases} \frac{xy}{x^2 + y^2} & \text{if } (x, y) \neq (0, 0) \\ 0 & \text{if } (x, y) = (0, 0). \end{cases}$$

Show that  $f$  is not continuous at  $(0, 0)$ .

**Solution:** This function has the value 0 at  $(0, 0)$ , but every disc centered at  $(0, 0)$  contains points of the form  $(x, x)$  with  $x \neq 0$  and, at such a point,  $f$  has the value  $1/2$ . So the condition for continuity at  $(0, 0)$  will not be satisfied when  $\epsilon$  is  $1/2$  or less.

**Example 8.1.4.** Show that the function with domain  $\mathbb{R}^2$  defined by

$$f(x, y) = \begin{cases} \frac{xy}{\sqrt{x^2 + y^2}} & \text{if } (x, y) \neq (0, 0) \\ 0 & \text{if } (x, y) = (0, 0) \end{cases}$$

is continuous at  $(0, 0)$ .

**Solution:** Since  $(x+y)^2 \geq 0$  and  $(x-y)^2 \geq 0$ , it follows that  $-2xy \leq x^2 + y^2$  and  $2xy \leq x^2 + y^2$ . Taken together, these two inequalities imply that

$$2|xy| \leq x^2 + y^2$$

On dividing by  $2\sqrt{x^2 + y^2}$  this becomes

$$|f(x, y) - f(0, 0)| = \left| \frac{xy}{\sqrt{x^2 + y^2}} \right| \leq \frac{1}{2} \sqrt{x^2 + y^2} = \frac{1}{2} \|(x, y) - (0, 0)\|.$$

Thus, given  $\epsilon > 0$ , if  $\delta = 2\epsilon$ , then

$$|f(x, y) - f(0, 0)| < \epsilon \quad \text{whenever} \quad \|(x, y) - (0, 0)\| < \delta.$$

We conclude that  $f$  is continuous at  $(0, 0)$ .

## Vector Valued Functions

The previous two examples involved real valued functions, We will also be concerned with functions with values in  $\mathbb{R}^q$  for some natural number  $q > 1$ . Given such a function  $F$  with domain  $D \subset \mathbb{R}^p$ , for each  $x \in D$  let  $f_j(x) = e_j \cdot F(x)$  be the  $j$ th component of the vector  $F(x) \in \mathbb{R}^q$ . Then each  $f_j$  is a real valued function on  $D$ . We will sometimes denote the function  $F$  by

$$F(x) = (f_1(x), f_2(x), \dots, f_q(x)).$$

The real valued function  $f_j$  is called the  $j$ th component function of  $F$ .

**Theorem 8.1.5.** A function  $F : D \rightarrow \mathbb{R}^q$  is continuous at a point  $a \in D$  if and only if each of its component functions is continuous at  $a$ .

*Proof.* It follows from Theorem 7.1.13 that, for each  $k$  and each  $x \in D$ ,

$$|f_k(x) - f_k(a)| \leq \|F(x) - F(a)\| \leq \sum_{j=1}^q |f_j(x) - f_j(a)|.$$

Given  $\epsilon > 0$ , It follows from the first inequality that if  $\|F(x) - F(a)\| < \epsilon$ , then also  $|f_k(x) - f_k(a)| < \epsilon$  for each  $k$ . Hence, if  $F$  is continuous at  $x_0$ , then so is each  $f_k$ . It follows from the second inequality that if  $|f_j(x) - f_j(a)| < \epsilon/q$  for each  $j$ , then  $\|F(x) - F(a)\| < \epsilon$ . This implies that if each  $f_j$  is continuous at  $a$ , then so is  $F$ .  $\square$

## Sequences and Continuity

Recall that Theorem 3.1.5 says that a function  $f$  of one variable is continuous at a point  $a$  of its domain  $D$  if and only if it takes sequences in  $D$  which converge to  $a$  to sequences which converge to  $f(a)$ . The same theorem is true of functions of several variables, in fact, it is true of any function from one metric space to another. The proof is also the same and we won't repeat it.

**Theorem 8.1.6.** *Let  $D$  be a subset of  $\mathbb{R}^p$ ,  $a \in D$ , and  $F : D \rightarrow \mathbb{R}^q$  a transformation. Then  $F$  is continuous at  $a$  if and only if, whenever  $\{x_n\}$  is a sequence in  $D$  which converges to  $a$ , then the sequence  $\{F(x_n)\}$  converges to  $F(a)$ .*

If  $F$  and  $G$  are two functions with domain  $D \subset \mathbb{R}^p$  and with values in  $\mathbb{R}^q$  and if  $h$  is a real valued function with domain  $D$ , then we can define new functions,  $hF$ ,  $F + G$ , and  $F \cdot G$  by

$$\begin{aligned} (hF)(x) &= h(x)F(x), \\ (F + G)(x) &= F(x) + G(x), \\ (F \cdot G)(x) &= F(x) \cdot G(x). \end{aligned} \tag{8.1.1}$$

Theorems 7.2.12 and 8.1.6 combine to prove the following theorem. The details are left to the exercises.

**Theorem 8.1.7.** *With  $F$ ,  $G$ ,  $h$ , and  $D$  as above, if  $F$ ,  $G$ , and  $h$  are continuous at  $a \in D$ , then so are  $hF$ ,  $F + G$ , and  $F \cdot G$ .*

## Composition of Functions

If  $G : D \rightarrow \mathbb{R}^p$  is a function with domain  $D \subset \mathbb{R}^d$  and  $F : E \rightarrow \mathbb{R}^q$  is a function with domain  $E \subset \mathbb{R}^p$ , then  $F(G(x))$  is defined as long as  $x \in D$  and  $G(x) \in E$ . Thus,

$$(F \circ G)(x) = F(G(x))$$

defines a function with domain  $D \cap G^{-1}(E)$  and with values in  $\mathbb{R}^q$ . This is the *composition* of the function  $G$  with the function  $F$ .

The following theorem follows immediately from two applications of Theorem 8.1.6. The details are left to the exercises.

**Theorem 8.1.8.** *With  $F$  and  $G$  as above, if  $a \in D \cap G^{-1}(E)$ ,  $G$  is continuous at  $a$ , and  $F$  is continuous at  $G(a)$ , then  $F \circ G$  is continuous at  $a$ .*

## Limits

Whether or not a function  $F$  is defined at a point  $a \in \mathbb{R}^p$ , it may have a limit as  $x$  approaches  $a$ . In order for this concept to make sense, it must be the case that there are points of the domain of  $F$  which are arbitrarily close but not equal to  $a$ .

If  $D$  is a subset of  $\mathbb{R}^p$  and  $a \in \mathbb{R}^p$ , then we will say that  $a$  is a *limit point* of  $D$  if every neighborhood of  $a$  contains points of  $D$  different from  $a$  (note that  $a$  may or may not be in  $D$ ).

**Definition 8.1.9.** If  $D \subset \mathbb{R}^p$ ,  $a$  is a limit point of  $D$ , and  $F : D \rightarrow \mathbb{R}^q$  is a function with domain  $D$ , then we will say that the limit of  $F$  as  $x$  approaches  $a$  is  $b$  if, for each  $\epsilon > 0$ , there is a  $\delta > 0$  such that

$$\|F(x) - b\| < \epsilon \quad \text{whenever} \quad x \in D \quad \text{and} \quad 0 < \|x - a\| < \delta.$$

In this case, we write  $\lim_{x \rightarrow a} F(x) = b$ .

If we compare this definition with the definition of continuity at  $a$  (Definition 8.1.1), we see that a function  $F : D \rightarrow \mathbb{R}^q$  is continuous at a point  $a \in D$  which is a limit point of  $D$  if and only if  $\lim_{x \rightarrow a} F(x) = F(a)$ .

On the other hand, if  $a \in D$  but  $a$  is not a limit point of  $D$ , then a function  $F$ , with domain  $D$  is automatically continuous at  $a$  (since, for small enough  $\delta$ , there are no points  $x \in D$  with  $\|x - a\| < \delta$  other than  $x = a$ ), but the limit of  $F$  as  $x$  approaches  $a$  is not defined. A point of  $D$  which is not a limit point of  $D$  is called an *isolated point* of  $D$ . For example, the set  $D = B_1((0,0)) \cup \{(1,1)\}$  is a subset of  $\mathbb{R}^2$  with  $(1,1)$  as an isolated point.

Note that Examples 8.1.3 and 8.1.4 show that

$$\lim_{(x,y) \rightarrow (0,0)} \frac{xy}{\sqrt{x^2 + y^2}} = 0,$$

while

$$\lim_{(x,y) \rightarrow (0,0)} \frac{xy}{x^2 + y^2}$$

does not exist. In fact, this function has limit

$$\frac{a}{1 + a^2}$$

as  $(x, y)$  approaches  $(0, 0)$  along the line  $y = ax$ . Since the function approaches different numbers as  $(x, y)$  approaches  $(0, 0)$  from different directions, the limit does not exist.

## Curves and Surfaces

A continuous function  $\gamma : I \rightarrow \mathbb{R}^q$ , where  $I$  is an interval in  $\mathbb{R}$ , is called a *parameterized curve* with parameter interval  $I$ . The variable  $t$  in  $\gamma(t)$  is called the *parameter* for the curve. Intuitively, as  $t$  ranges through the parameter interval,  $\gamma(t)$  traces out something like a curved line in  $\mathbb{R}^q$ .

If the parameter interval  $I$  is a closed bounded interval  $[a, b]$  with  $\gamma(a) = x$  and  $\gamma(b) = y$ , then  $\gamma$  is called a curve in  $\mathbb{R}^q$  joining  $x$  to  $y$ . The points  $x$  and  $y$  are called the *endpoints* of the curve. If  $x = y$ , then  $\gamma$  is called a closed curve.

**Example 8.1.10.** Give examples of a closed curve, a curve with endpoints which is not closed, and a curve with no endpoints.

**Solution:** The curve  $\gamma(t) = (\cos t, \sin t)$ ,  $t \in [0, 2\pi]$ , is a closed curve in  $\mathbb{R}^2$ . It is closed because  $\gamma(0) = (1, 0) = \gamma(2\pi)$ .

The curve  $\gamma(t) = (t^2, t^3)$ ,  $t \in [0, 1]$ , is a curve joining  $x = (0, 0)$  and  $y = (1, 1)$ . It has these points as endpoints. It is not closed, since the endpoints are not the same.

The curve  $\gamma(t) = (t \cos t, t \sin t, t)$ ,  $t \in (-\infty, \infty)$  is a spiral curve in  $\mathbb{R}^3$  with no endpoints.

Generally, a curve is a one dimensional object, but there are exceptions. A curve may be *degenerate* – that is,  $\gamma(t)$  may be a constant vector in  $\mathbb{R}^q$ . Then the image of  $\gamma$  is a single point, which is a zero dimensional object.

A *parameterized surface* in  $\mathbb{R}^q$  ( $q \geq 2$ ) is a continuous function  $F : A \rightarrow \mathbb{R}^q$ , where  $A$  is an open subset of  $\mathbb{R}^2$  or an open subset of  $\mathbb{R}^2$  together with all or part of the boundary of this open subset.

**Example 8.1.11.** Give three examples of parameterized surfaces.

**Solution:** The image of the surface

$$F(\theta, \phi) = (\cos \theta \cos \phi, \sin \theta \cos \phi, \sin \phi) \quad \text{with} \quad \theta \in [0, 2\pi), \phi \in [0, \pi]$$

is the sphere of radius 1 centered at the origin. The parameter set  $A$  in this case is the rectangle  $[0, 2\pi) \times [0, \pi]$ . The parameterization is the one given by expressing the sphere in spherical coordinates. Note that this sphere is just  $\overline{B_1(0)} \setminus B_1(0)$  and, hence, is a closed set (Exercise 7.3.6) even though its parameter set is not closed.

The closed upper half of the above sphere may be parameterized as above but with parameter set  $[0, 2\pi) \times [0, \pi/2]$  or it may be parameterized by

$$G(x, y) = (x, y, \sqrt{1 - x^2 - y^2}) \quad \text{with} \quad x^2 + y^2 \leq 1.$$

Here, the set  $A$  is the closed disc of radius 1 centered at the origin in  $\mathbb{R}^2$ .

If we change the parameter set for  $G$  in the above example to the open disc of radius 1 centered at 0, then we obtain a surface which is not a closed set – the upper half of the unit sphere not including the circle  $\{(x, y, z) : x^2 + y^2 = 1, z = 0\}$ .

Generally, the image of a parameterized surface is a two dimensional object, but there are exceptions. A surface may be *degenerate*. The parameter function  $F$  could have image contained in a set of dimension less than 2 – it could be a point, or a curve. For example, the image of

$$F(u, v) = (\cos(u + v), \sin(u + v), u + v) \quad \text{with } (u, v) \in \mathbb{R}^2$$

is actually the spiral curve  $(\cos t, \sin t, t)$ , as we can see by making the substitution  $t = u + v$ .

Conditions that guarantee that a curve or surface is not degenerate will be obtained in the next chapter.

### Exercise Set 8.1

1. Consider the function  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  defined by

$$f(x, y) = \begin{cases} \frac{xy^2}{x^2 + y^2} & \text{if } (x, y) \neq (0, 0) \\ 0 & \text{if } (x, y) = (0, 0). \end{cases}$$

Is this function continuous at  $(0, 0)$ ? Justify your answer.

2. Give a simple reason why the function  $\gamma : \mathbb{R} \rightarrow \mathbb{R}^4$  defined by

$$\gamma(t) = (t, \sin t, e^t, t^2)$$

is continuous on  $\mathbb{R}$ .

3. Does the function  $f : \mathbb{R}^2 \setminus \{(0, 0)\} \rightarrow \mathbb{R}$ , defined by

$$f(x, y) = \frac{x}{\sqrt{x^2 + y^2}},$$

have a limit as  $(x, y)$  approaches  $(0, 0)$ . Justify your answer.

4. Consider the function  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  defined by

$$f(x, y) = \begin{cases} xy & \text{if } xy > 0 \\ 0 & \text{if } xy \leq 0. \end{cases}$$

At which points of  $\mathbb{R}^2$  is this function continuous?

5. For the function  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  defined by

$$f(x, y) = \frac{x^2 y}{x^4 + y^2}$$

Show that  $f$  has limit 0 as  $(x, y) \rightarrow (0, 0)$  along any straight line through the origin, but it does not have a limit as  $(x, y) \rightarrow (0, 0)$  in  $\mathbb{R}^2$ .

6. Consider the function  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  defined by

$$f(x, y) = \begin{cases} \frac{y^2 - x^2 y}{|y - x^2|} & \text{if } y \neq x^2 \\ 0 & \text{if } y = x^2. \end{cases}$$

At which points of  $\mathbb{R}^2$  is this function continuous?

7. Prove Theorem 8.1.7.

8. Prove Theorem 8.1.8.

9. Prove that  $a$  is a limit point of a set  $D \subset \mathbb{R}^p$  if and only if there is a sequence of points in  $D$  but not equal to  $a$  which converges to  $a$ .

10. Let  $D$  be a subset of  $\mathbb{R}^p$  and  $F : D \rightarrow \mathbb{R}^q$  a function. If  $a$  is a limit point of  $D$ , prove that  $\lim_{x \rightarrow a} F(x) = b$  if and only if  $\lim_{n \rightarrow \infty} F(x_n) = b$  whenever  $\{x_n\}$  is a sequence in  $D$  which converges to  $a$ .

11. Let  $F : D \rightarrow \mathbb{R}^q$  be a transformation with domain  $D \subset \mathbb{R}^p$  and let  $a$  be a limit point of  $D$ . Prove that if  $\{F(x_n)\}$  converges whenever  $\{x_n\}$  is a sequence in  $D$  which converges to  $a$ , then  $\lim_{x \rightarrow a} F(x)$  exists.

12. Let  $B_1(0)$  be the open unit ball in  $\mathbb{R}^2$ . Does every continuous function  $f : B_1(0) \rightarrow \mathbb{R}$  take Cauchy sequences to Cauchy sequences?

13. Let  $\overline{B}_1(0)$  be the closed unit ball in  $\mathbb{R}^2$ . Does every continuous function  $f : \overline{B}_1(0) \rightarrow \mathbb{R}$  take Cauchy sequences to Cauchy sequences?

14. Find a parameterized curve  $\gamma(t)$  in  $\mathbb{R}^2$ , with parameter interval  $[0, \infty)$ , that begins at  $(1, 0)$ , spirals inward in the counterclockwise direction, and approaches  $(0, 0)$  as  $t \rightarrow \infty$ .

15. Find a parameterization of the cylindrical surface in  $\mathbb{R}^3$  defined by the equation  $x^2 + y^2 = 1$  ( $z$  is unrestricted). That is, find a continuous function  $F : A \rightarrow \mathbb{R}^3$  with  $A \subset \mathbb{R}^2$ , such that  $F$  has the cylinder as image.

## 8.2 Properties of Continuous Functions

The theme of this section is that continuous functions are the functions that behave well with respect to topological properties of sets.

### Continuity and Open and Closed Sets

Recall that if  $D$  is a subset of  $\mathbb{R}^p$ , then a *relatively open* subset of  $D$  is a set of the form  $U \cap D$ , where  $U$  is open in  $\mathbb{R}^p$ . The relatively open subsets of  $D$  are the open subsets of  $D$  considered as a metric space by itself (rather than a subset of  $\mathbb{R}^p$ ). Relatively closed sets are defined analogously.

**Theorem 8.2.1.** *If  $D \subset \mathbb{R}^p$  and  $F : D \rightarrow \mathbb{R}^q$  is a function, then  $F$  is continuous on  $D$  if and only if  $F^{-1}(U)$  is a relatively open subset of  $D$  whenever  $U$  is an open subset of  $\mathbb{R}^q$ . Equivalently,  $F$  is continuous if and only if  $F^{-1}(A)$  is a relatively closed subset of  $D$  whenever  $A$  is a closed subset of  $\mathbb{R}^q$ .*

*Proof.* Suppose  $F$  is continuous and  $U$  is an open subset of  $\mathbb{R}^q$ . If  $a \in F^{-1}(U)$ , then  $b = F(a) \in U$ . Since  $U$  is open, there is an  $\epsilon > 0$  such that  $B_\epsilon(b) \subset U$ . Since  $F$  is continuous on  $D$ , there is a  $\delta > 0$  such that

$$\|F(x) - F(a)\| < \epsilon \quad \text{whenever } x \in D \quad \text{and} \quad \|x - a\| < \delta.$$

This implies that  $F(B_\delta(a) \cap D) \subset B_\epsilon(b) \subset U$ , and, hence, that

$$B_\delta(a) \cap D \subset F^{-1}(U).$$

Since we can do this at each  $a \in F^{-1}(U)$ , we conclude that  $F^{-1}(U)$  is the intersection of  $D$  with the union of the resulting collection of open balls  $B_\delta(a)$ . Hence, it is relatively open in  $D$ .

On the other hand, suppose  $F^{-1}(U)$  is relatively open in  $D$  for each open set  $U$  in  $\mathbb{R}^q$ . In particular, this implies that if  $a \in D$ ,  $b = F(a)$ , and  $\epsilon > 0$ , then the set  $F^{-1}(B_\epsilon(b))$  is relatively open in  $D$ . Thus,

$$F^{-1}(B_\epsilon(b)) = D \cap V$$

for some open set  $V \subset \mathbb{R}^p$ . Since  $a \in V$  and  $V$  is open, there is a  $\delta > 0$  such that  $B_\delta(a) \subset V$ . Then  $x \in D$  and  $\|x - a\| < \delta$  implies  $x \in V \cap D = F^{-1}(B_\epsilon(b))$ . This means that

$$\|F(x) - F(a)\| < \epsilon \quad \text{whenever } x \in D \quad \text{and} \quad \|x - a\| < \delta.$$

Hence,  $F$  is continuous at  $a$ . Since this is true for all points  $a \in D$ , we conclude that  $F$  is continuous on  $D$ .

The analogous result for closed sets follows from the above by taking complements and using the fact that a subset of  $D$  is relatively closed if and only if it is the complement in  $D$  of a set which is relatively open. The details are left to the exercises.  $\square$

If  $D$  is open, then the relatively open subsets of  $D$  are just the open subsets of  $D$ . Hence, we have the following corollary of the above theorem.

**Corollary 8.2.2.** *If  $D \subset \mathbb{R}^p$  is open and  $F : D \rightarrow \mathbb{R}^q$  is a function, then  $F$  is continuous on  $D$  if and only if  $F^{-1}(U)$  is open for every open set  $U \subset \mathbb{R}^q$ .*

## Continuity and Compactness

The proof of the following theorem is very simple, but it has a lot of very useful consequences.

**Theorem 8.2.3.** *If  $K$  is a compact subset of  $\mathbb{R}^p$  and  $F : K \rightarrow \mathbb{R}^q$  is a continuous function, then  $F(K)$  is a compact subset of  $\mathbb{R}^q$ .*

*Proof.* Let  $\mathcal{U}$  be an open cover of  $F(K)$  and let  $\mathcal{V}$  be the collection of all open subsets  $V \subset \mathbb{R}^p$  such that  $V \cap K = F^{-1}(U)$  for some  $U \in \mathcal{U}$ . There is at least one such  $V$  for each  $U \in \mathcal{U}$  since  $F^{-1}(U)$  is relatively open in  $K$  by the previous theorem.

Since  $\mathcal{U}$  is a cover of  $F(K)$ ,  $\mathcal{V}$  is an open cover of  $K$ . Since  $K$  is compact, there is a finite subcollection  $\{V_j\}_{j=1}^n$  of  $\mathcal{V}$  which also covers  $K$ . For each  $V_j$  we may choose a  $U_j \in \mathcal{U}$  such that  $V_j \cap K = F^{-1}(U_j)$ .

If  $y \in F(K)$ , then  $y = F(x)$  for some  $x \in K$ . This  $x$  belongs to  $V_j \cap K$  for some  $j$  because  $\{V_j\}_{j=1}^n$  is a cover of  $K$ . Then  $y \in U_j$ . This proves that the collection  $\{U_j\}_{j=1}^n$  is a cover of  $F(K)$ . It is, in fact, a finite subcover of  $\mathcal{U}$ . Since we can do this for every open cover of  $F(K)$ , we have proved that  $F(K)$  is compact.  $\square$

A function  $F : D \rightarrow \mathbb{R}^q$  is said to be *bounded* on  $D$  if there is a number  $M$  such that

$$\|F(x)\| \leq M \quad \text{for all } x \in D.$$

That is,  $F$  is bounded on  $D$  if the set of non-negative numbers  $\{\|F(x)\| : x \in D\}$  is bounded above. The least upper bound of this set is denoted  $\sup_D \|F(x)\|$ . It may or may not be a member of the set – that is, there may or may not be a point  $x_0 \in D$  such that  $\|F(x_0)\| = \sup_D \|F(x)\|$ . If there is such a point  $x_0$ , then we say that  $\|F(x)\|$  assumes a maximum value on  $D$ .

A compact set contains points of maximal norm and points of minimal norm (Exercise 7.4.4). Combining this with the previous theorem yields the following:

**Theorem 8.2.4.** *If  $K \subset \mathbb{R}^p$  is compact and  $F : K \rightarrow \mathbb{R}^q$  is continuous, then  $F$  is bounded on  $K$  and  $\|F(x)\|$  assumes a maximum value on  $K$ .*

*Proof.* By the previous theorem,  $F(K)$  is compact and, hence, bounded. Furthermore, it contains a point of maximum norm by Exercise 7.4.4. This point is in  $F(K)$  and so it the form  $F(x_0)$  for some  $x_0 \in K$ .  $\square$

**Corollary 8.2.5.** *If  $K \subset \mathbb{R}^p$  is compact and  $f : K \rightarrow \mathbb{R}$  is a continuous real valued function on  $K$ , then  $f$  assumes a maximal value and a minimal value on  $K$ .*

*Proof.* It follows from the previous theorem that  $\{|f(x)| : x \in K\}$  is bounded above by some number  $M$ . Then the function  $g(x) = f(x) + M$  is a non-negative function and so  $|g(x)| = g(x)$ . By the previous theorem, there is a point  $x_0 \in K$  with

$$g(x) \leq g(x_0) \quad \text{for all } x \in K.$$

Since  $f(x) = g(x) - M$ , it follows that  $x_0$  is a point at which  $f$  achieves its maximal value.

Since the above argument applies equally well to  $-f(x)$ , and, since a maximum for  $-f(x)$  on  $K$  will be the negative of a minimum for  $f(x)$  on  $K$ , it follows that  $f(x)$  has a minimum value on  $K$  as well.  $\square$

**Example 8.2.6.** Let  $K$  be a compact subset of  $\mathbb{R}^p$ . Show that  $f : K \rightarrow \mathbb{R}$  is a real valued continuous function on  $K$  which is strictly positive at each point of  $K$ , then there is a number  $\delta > 0$  such that  $f(x) \geq \delta$  for all  $x \in K$ .

**Solution:** By Corollary 8.2.5, the function  $f$  has a minimum value  $\delta$  on  $K$ . This minimum value cannot be 0, since  $f$  is positive at all points of  $K$ . Thus,  $\delta > 0$  and  $f(x) \geq \delta$  for all  $x \in K$ .

## Continuity and Connectedness

Continuous functions also take connected sets to connected sets.

**Theorem 8.2.7.** *If  $D \subset \mathbb{R}^p$  is connected and  $F : D \rightarrow \mathbb{R}^q$  is continuous, then  $F(D)$  is also connected.*

*Proof.* Suppose  $U$  and  $V$  are open subsets of  $\mathbb{R}^q$  such that  $U \cap V = \emptyset$  and  $F(D) \subset U \cup V$ . Then  $F^{-1}(U)$  and  $F^{-1}(V)$  are relatively open subsets of  $D$ ,  $F^{-1}(U) \cap F^{-1}(V) = \emptyset$ , and  $D \subset F^{-1}(U) \cup F^{-1}(V)$ . Thus, one of the sets  $F^{-1}(U) \cap D$  and  $F^{-1}(V) \cap D$  must be empty since, otherwise, they would separate  $D$ . However, if  $F^{-1}(U) \cap D = \emptyset$ , then  $U \cap F(D) = \emptyset$  and a similar statement holds for  $V$ . Thus, either  $U$  or  $V$  has empty intersection with  $F(D)$  which implies that the two sets do not separate  $F(D)$ . Hence,  $F(D)$  is connected.  $\square$

The following is the several variable version of the Intermediate Value Theorem, since it says that if a continuous real valued function on a connected set takes on two values, it also takes on every value in between the two.

**Corollary 8.2.8.** *If  $D \subset \mathbb{R}^p$  is connected and  $f : \mathbb{R}^p \rightarrow \mathbb{R}$  is a continuous function, then  $f(D)$  is an interval.*

*Proof.* By the previous theorem,  $f(D)$  is a connected subset of the line  $\mathbb{R}$ . By Theorem 7.5.4 the only such sets are intervals.  $\square$

Now suppose  $E$  is a subset of  $\mathbb{R}^d$  and  $\gamma : I \rightarrow E$  is a parameterized curve with parameter interval  $I = [a, b]$ . Since  $I$  is connected by Theorem 7.5.4, its image  $\gamma(I)$  is a connected subset of  $E$ . Thus, if  $x = \gamma(a)$  and  $y = \gamma(b)$ , then  $x$  and  $y$  must be in the same component of  $E$ . Thus, we have proved the following.

**Theorem 8.2.9.** *If  $E$  is a subset of  $\mathbb{R}^d$  and  $x$  and  $y$  are points of  $E$  that may be joined by a curve in  $E$ , then  $x$  and  $y$  are in the same connected component of  $E$ . If each pair of points of  $E$  may be joined by a curve in  $E$ , then  $E$  is connected.*

**Example 8.2.10.** Show that the unit circle  $T$  (the set of points  $(x, y) \in \mathbb{R}^2$  with  $x^2 + y^2 = 1$ ) is connected.

**Solution:** Each point on the circle  $T$  is of the form  $(\cos t, \sin t)$ . Each pair of such points  $(\cos a, \sin a)$  and  $(\cos b, \sin b)$  with  $a < b$ , are joined by the curve

$$\gamma(t) = (\cos t, \sin t) \quad t \in [a, b]$$

which lies in the circle. Hence, the circle  $T$  is connected.

### Uniform Continuity

**Definition 8.2.11.** Let  $D$  be a subset of  $\mathbb{R}^p$  and  $F : D \rightarrow \mathbb{R}^q$  a function. Then  $F$  is said to be *uniformly continuous* on  $D$  if for each  $\epsilon > 0$  there is a  $\delta > 0$  such that

$$\|F(x) - F(y)\| < \epsilon \quad \text{whenever} \quad x, y \in D \quad \text{and} \quad \|x - y\| < \delta.$$

As with uniform continuity for functions of one variable, discussed in Section 3.3, the point here is that the choice of  $\delta$  does not depend on  $x$  or  $y$ .

Uniform continuity is an important concept and it will play a key role in our proof of the existence of the Riemann integral of a function of several variables.

We proved in Theorem 3.3.4 that a continuous function on closed, bounded interval is uniformly continuous. The analogous theorem holds for functions of several variables, but compact sets replace closed, bounded intervals.

**Theorem 8.2.12.** *If  $K$  is a compact subset of  $\mathbb{R}^p$  and  $F : K \rightarrow \mathbb{R}^q$  is continuous on  $K$ , then  $F$  is uniformly continuous on  $K$ .*

*Proof.* Since  $F$  is continuous on  $K$ , given  $\epsilon > 0$  we may choose for each  $x \in K$  a number  $\delta(x) > 0$  such that

$$\|F(y) - F(x)\| < \epsilon/2 \quad \text{whenever} \quad y \in K \quad \text{and} \quad \|y - x\| < \delta(x). \quad (8.2.1)$$

We set  $\rho(x) = \delta(x)/2$ . Then  $\rho(x)$  is a positive valued function defined on  $K$ , just as in Example 7.4.9. In that example, we showed that a consequence of the compactness of  $K$  is that there is a finite set of points  $\{x_1, x_2, \dots, x_n\}$  such that  $K$  is contained in the union of the balls  $B_{\rho(x_j)}(x_j)$  for  $j = 1, \dots, n$ .

We set  $\rho = \min\{\rho(x_j) : j = 1, \dots, n\}$ . Then given any two points  $x, y \in K$  with  $\|x - y\| < \rho$ ,  $x$  must be in  $B_{\rho(x_j)}(x_j)$  for some  $j$ . This implies that  $\|x - x_j\| < \rho(x_j) < \delta(x_j)$  and

$$\|y - x_j\| \leq \|y - x\| + \|x - x_j\| < \rho + \rho(x_j) \leq 2\rho(x_j) = \delta(x_j).$$

Since both  $x$  and  $y$  are within  $\delta(x_j)$  of  $x_j$ , it follows from (8.2.1) that

$$\|F(x) - F(y)\| \leq \|F(x) - F(x_j)\| + \|F(x_j) - F(y)\| < \epsilon/2 + \epsilon/2 = \epsilon.$$

Hence,  $F$  is uniformly continuous on  $K$ . □

In Theorem 3.3.6 we showed that a function is uniformly continuous on a bounded interval if and only if it has a continuous extension to the closure of the interval. The analogous theorem holds for functions from  $\mathbb{R}^p$  to  $\mathbb{R}^q$ .

**Theorem 8.2.13.** *If  $D \subset \mathbb{R}^p$  is a bounded set and  $F : D \rightarrow \mathbb{R}^q$  is a function, then  $F$  is uniformly continuous on  $D$  if and only if  $F$  can be extended to a continuous function  $\hat{F} : \overline{D} \rightarrow \mathbb{R}^q$ .*

*Proof.* Note that, since  $D$  is bounded,  $\overline{D}$  is compact. Thus, if  $F$  has an extension to a continuous function  $\hat{F} : \overline{D} \rightarrow \mathbb{R}^q$ , then  $\hat{F}$  is uniformly continuous on  $\overline{D}$ , by the previous theorem. Then  $\hat{F}$  is also uniformly continuous on the smaller set  $D$ . But  $\hat{F} = F$  on  $D$ , and so  $F$  is uniformly continuous on  $D$ .

Conversely, suppose  $F$  is uniformly continuous on  $D$ . Then  $\{F(x_n)\}$  is a Cauchy sequence in  $\mathbb{R}^q$  whenever  $\{x_n\}$  is a Cauchy sequence in  $D$  (Exercise 8.2.11). If  $x \in \overline{D}$ , then there is a sequence  $\{x_n\}$  in  $D$  that converges to  $x$  (Theorem 7.3.10). Such a sequence is necessarily Cauchy and so  $\{F(x_n)\}$  is also Cauchy. But Cauchy sequences in  $\mathbb{R}^q$  converge by Theorem 7.2.16.

If  $\{y_n\}$  is another sequence in  $D$  which converges to  $x$ , then we may construct a third sequence  $\{z_n\}$  converging to  $x$  by intertwining the sequences  $\{x_n\}$  and  $\{y_n\}$  – that is, let  $z_{2n} = y_n$  and  $z_{2n-1} = x_n$ . Then,  $\{z_n\}$  not only converges to  $x$ , it has both  $\{x_n\}$  and  $\{y_n\}$  as subsequences. By the above argument, the sequence  $\{F(z_n)\}$  must converge to a point  $u \in \mathbb{R}^q$ . Both subsequences  $\{F(x_n)\}$  and  $\{F(y_n)\}$  must then converge to the same point  $u$ . Thus, we have proved that no matter what sequence  $\{x_n\}$  converging to  $x$  we choose, the limit of the sequence  $\{F(x_n)\}$  is the same. Therefore, it makes sense to define an extension  $\hat{F}$  of  $F$  to  $\overline{D}$  by setting

$$\hat{F}(x) = \lim F(x_n)$$

for any sequence  $\{x_n\}$  in  $D$  converging to  $x$ . The resulting function is obviously equal to  $F$  on  $D$ , since we may just choose  $x_n = x$  for all  $n$  if  $x \in D$ .

We now have an extension  $\hat{F}$  of  $F$  to  $\overline{D}$ . It remains to prove that it is continuous on  $\overline{D}$ . We will do this by applying Theorem 8.1.6. If  $\{x_n\}$  is a sequence in  $\overline{D}$  which converges to  $x \in \overline{D}$ , we may choose for each  $n$  a point  $y_n \in D$  such that  $\|x_n - y_n\| < 1/n$  and  $\|F(y_n) - \hat{F}(x_n)\| < 1/n$ . Then

$$\|x - y_n\| \leq \|x - x_n\| + \|x_n - y_n\| < \|x - x_n\| + 1/n.$$

Since  $\|x - x_n\| \rightarrow 0$  and  $1/n \rightarrow 0$ , it follows that  $y_n \rightarrow x$  and, hence,  $F(y_n) \rightarrow \hat{F}(x)$  by our definition of  $\hat{F}$ . However, it also follows that  $\hat{F}(x_n) \rightarrow \hat{F}(x)$  since,

$$\|\hat{F}(x) - \hat{F}(x_n)\| \leq \|\hat{F}(x) - F(y_n)\| + \|F(y_n) - \hat{F}(x_n)\|,$$

and both  $\|F(y_n) - \hat{F}(x_n)\|$  and  $\|\hat{F}(x) - F(y_n)\|$  converge to 0.

Since  $\hat{F}(x_n) \rightarrow \hat{F}(x)$  whenever  $\{x_n\}$  is a sequence in  $\overline{D}$  converging to  $x \in \overline{D}$ , the function  $\hat{F}$  is continuous on  $\overline{D}$  by Theorem 8.1.6.  $\square$

### Exercise Set 8.2

1. If  $A = \{(x, y) \in \mathbb{R}^2 : 0 \leq x \leq 1, 0 \leq y \leq 1\}$ . Which of the following sets cannot be the image of the set  $A$  under a continuous function  $F : A \rightarrow \mathbb{R}^2$ ? Justify your answers.

- (a)  $\overline{B}_2(0, 0)$ ;
- (b)  $B_1(0)$ ;
- (c)  $\{(x, y) \in \mathbb{R}^2 : 0 \leq x \leq 1, 0 \leq y\}$ ;

- (d)  $\overline{B}_1(0, 0) \cup \overline{B}_1(3, 0)$ .  
 (e)  $\{(t, t) \in \mathbb{R}^2 : t \in \mathbb{R}; 0 \leq t \leq 1\}$ .

2. Finish the proof of Theorem 8.2.1, by proving that a function is continuous if and only if the inverse image of each closed set is relatively closed. Hint: you may use the first part of the theorem (that a function is continuous if and only if the inverse image of each open set is relatively open).
3. If  $K$  is a compact, connected subset of  $\mathbb{R}^p$  and  $f : K \rightarrow \mathbb{R}$  is a continuous function, what can you say about  $f(K)$ ?
4. If  $F : \mathbb{R}^p \rightarrow \mathbb{R}^q$  is continuous and  $A$  is a bounded subset of  $\mathbb{R}^p$ , prove that  $\overline{F(A)} = F(\overline{A})$ . Is this necessarily true if  $A$  is not bounded?
5. The image of a compact set under a continuous function is compact, hence closed, by Theorem 8.2.3. Is the image of a closed set under a continuous function necessarily closed? Prove that it is or give an example where it is not.
6. Is the image of an open set under a continuous function necessarily an open set? Prove that it is or give an example where it is not.
7. Is the sphere  $\{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = 1\}$  connected? How do you know?
8. Prove that if  $f : T \rightarrow \mathbb{R}$  is a continuous real valued function on the unit circle  $T$ , then there is a pair of diametrically opposed points  $(x, y)$  and  $(-x, -y)$  on  $T$  at which  $f$  has the same value.
9. Find an example of a closed set  $A \subset \mathbb{R}^2$ , which is connected, but which contains two points that cannot be joined by a curve in  $A$ .
10. Is the function  $f : \mathbb{R}^2 \setminus \{(2, 0)\} \rightarrow \mathbb{R}$  defined by

$$f(x, y) = \frac{1}{(x-2)^2 + y^2}$$

uniformly continuous on  $B_1(0, 0)$ ? Is it uniformly continuous on  $B_2(0, 0)$ ? Justify your answers.

11. If  $D \subset \mathbb{R}^p$ , prove that if a function  $F : D \rightarrow \mathbb{R}^q$  is uniformly continuous on  $D$  then  $\{F(x_n)\}$  is a Cauchy sequence in  $\mathbb{R}^q$  whenever  $\{x_n\}$  is a Cauchy sequence in  $D$ .
12. Show that the converse of the statement in the previous exercise is not true in general, but it is true if the set  $D$  is bounded. That is, show that there exist a  $D$  and a continuous function  $F : D \rightarrow \mathbb{R}^q$  which is not uniformly continuous but which does take each Cauchy sequence in  $D$  to a Cauchy sequence in  $\mathbb{R}^d$ . However, show there are no such functions if  $D$  is bounded.

13. Does uniform continuity make sense for a function from one metric space to another? If so, how would you define it?

### 8.3 Sequences of Functions

Uniform convergence of sequences of functions will play the same role in functions of several variables that it did in earlier chapters on functions of a single variable. It preserves continuity and allows the limit to be taken inside an integral.

The results of Section 3.4 on uniform convergence hold in the several variable context and have almost the same proofs.

#### Uniform convergence

**Definition 8.3.1.** Let  $\{F_n\}$  be a sequence of functions from  $D$  to  $\mathbb{R}^q$ , where  $D \subset \mathbb{R}^p$ . We say this sequence converges pointwise to  $F : D \rightarrow \mathbb{R}^q$  on  $D$  if the sequence  $\{F_n(x)\}$  converges to  $F(x)$  for each  $x \in D$ .

We say  $\{F_n\}$  converges *uniformly* to  $F : D \rightarrow \mathbb{R}^q$  on  $D$  if, for each  $\epsilon > 0$ , there is an  $N$  such that

$$\|F(x) - F_n(x)\| < \epsilon \quad \text{whenever } x \in D \text{ and } n \geq N.$$

The difference between pointwise and uniform convergence is that, in the latter, the choice of  $N$  must be independent of  $x$ .

The following test for uniform convergence is the several variable analogue of Theorem 3.4.6. The proof is simple and is left to the exercises.

**Theorem 8.3.2.** Let  $F$  be a function and  $\{F_n\}$  a sequence of functions defined on a set  $D \subset \mathbb{R}^p$  and having values in  $\mathbb{R}^q$ . If there is a sequence of non-negative numbers  $\{b_n\}$ , such that  $b_n \rightarrow 0$ , and

$$\|F(x) - F_n(x)\| \leq b_n \quad \text{for all } x \in D,$$

then  $\{F_n\}$  converges uniformly to  $F$  on  $D$ .

**Example 8.3.3.** Examine the convergence of the sequence  $\{(x^2 + y^2)^n\}$  on the closed disc  $\overline{B}_r(0, 0)$  in  $\mathbb{R}^2$  for each  $r \leq 1$ .

**Solution:** Note that  $x^2 + y^2 \leq r^2$  on  $\overline{B}_r(0, 0)$ . Thus,

$$|(x^2 + y^2)^n| \leq r^{2n} \quad \text{on } \overline{B}_r(0, 0).$$

If  $r < 1$ , then  $r^{2n} \rightarrow 0$  and, hence,  $\{(x^2 + y^2)^n\}$  converges uniformly to 0 on  $\overline{B}_r(0, 0)$  by the previous theorem.

On  $\overline{B}_1(0, 0)$ , the sequence  $\{(x^2 + y^2)^n\}$  converges to 0 if  $(x, y)$  is in the interior of the disc and to 1 if  $(x, y)$  is on the boundary of the disc. The limit function is not continuous on  $\overline{B}_1(0, 0)$  and, by the next theorem, this means the convergence is not uniform. Without using the next theorem, we can still

easily see that the convergence is not uniform – in fact, not uniform even on the smaller set  $B_1(0, 0)$ . Given an  $\epsilon$  with  $0 < \epsilon < 1$ , if  $(x, y) \in B_1(0, 0)$  and we set  $r = \|(x, y)\| < 1$ , then  $|(x^2 + y^2)^n| = r^{2n}$  and so

$$|(x^2 + y^2)^n| < \epsilon \quad (8.3.1)$$

if and only if  $r^{2n} < \epsilon$ , which holds if and only if

$$n > N_r = \frac{\ln \epsilon}{2 \ln r}.$$

Thus, an  $N$  with the property that (8.3.1) holds for all  $r < 1$  must be larger than  $N_r$  for all  $r < 1$ . There is no such  $N$ , since  $\lim_{r \rightarrow 1} N_r = \infty$ .

### Uniform Convergence and Continuity

One of the main reasons uniform convergence is important is the following theorem. Its proof is the same as the proof of the analogous theorem for real valued functions of a real variable (Theorem 3.4.4), and we will not repeat it.

**Theorem 8.3.4.** *If  $\{F_n\}$  is a sequence of continuous functions from a subset  $D$  of  $\mathbb{R}^p$  to  $\mathbb{R}^q$ , which converges uniformly on  $D$  to a function  $F$ , then  $F$  is also continuous on  $D$ .*

As we saw in example 8.3.3, a sequence of continuous functions which converges only pointwise may not converge to a continuous function.

**Example 8.3.5.** Define a sequence  $\{F_n\}$  of functions from the unit ball  $B_1(0, 0)$  in  $\mathbb{R}^2$  to  $\mathbb{R}^2$  by

$$F_n(x, y) = \left( \frac{x^2 - ny^2}{1 + ny^2}, \frac{nx}{1 + nx^2} \right).$$

Show that this sequence converges pointwise, but not uniformly on  $B_1(0, 0)$ .

**Solution:** Each of the functions  $F_n$  is continuous on  $B_1(0, 0)$ . The sequence clearly converges pointwise to the function  $F$  defined on  $B_1(0, 0)$  by

$$F(x) = \begin{cases} (-1, 1/x) & \text{if } x \neq 0, y \neq 0 \\ (-1, 0) & \text{if } x = 0, y \neq 0 \\ (x^2, 1/x) & \text{if } x \neq 0, y = 0 \\ (0, 0) & \text{if } x = 0, y = 0 \end{cases}$$

This function is not continuous on  $B_1(0, 0)$  – in fact, it is discontinuous at all points on the  $x$  and  $y$  axes – and so, by the previous theorem, the convergence of  $\{F_n\}$  to  $F$  cannot be uniform on  $B_1(0, 0)$ .

### Uniformly Cauchy Sequences

**Definition 8.3.6.** If  $D \subset \mathbb{R}^p$  and  $\{F_n\}$  is a sequence of functions from  $D$  to  $\mathbb{R}^q$ , then  $\{F_n\}$  is said to be *uniformly Cauchy* if, for each  $\epsilon > 0$ , there is an  $N$  such that

$$\|F_n(x) - F_m(x)\| < \epsilon \quad \text{whenever } x \in D \text{ and } n, m \geq N.$$

Another several variable analogue of a single variable theorem (Theorem 3.4.10) is the following. Since the proof of the single variable version was left to the exercises, we will actually prove this version.

**Theorem 8.3.7.** *If  $D \subset \mathbb{R}^p$ , a sequence of functions  $F_n : D \rightarrow \mathbb{R}^q$  is uniformly Cauchy if and only if it converges uniformly to some function  $F : D \rightarrow \mathbb{R}^q$ .*

*Proof.* If  $\{F_n\}$  converges uniformly on  $D$  to a function  $F$  and  $\epsilon > 0$ , then there is an  $N$  such that

$$\|F(x) - F_n(x)\| < \epsilon/2 \quad \text{whenever } x \in D, n \geq N.$$

Then

$$\|F_n(x) - F_m(x)\| \leq \|F_n(x) - F(x)\| + \|F(x) - F_m(x)\| < \epsilon/2 + \epsilon/2 = \epsilon$$

whenever  $x \in D$  and  $n, m \geq N$ . Thus,  $\{F_n\}$  is uniformly Cauchy.

On the other hand, if  $\{F_n\}$  is uniformly Cauchy, then for each  $x \in D$ ,  $\{F_n(x)\}$  is a Cauchy sequence of vectors in  $\mathbb{R}^q$  and, hence, converges to some vector  $F(x) \in \mathbb{R}^q$  by Theorem 7.2.16. That is,  $\{F_n\}$  converges pointwise to a function  $F : D \rightarrow \mathbb{R}^q$ . It remains to prove that the convergence is uniform.

Since the sequence is uniformly Cauchy, for each  $\epsilon > 0$  there is an  $N$  such that

$$\|F_n(x) - F_m(x)\| < \epsilon/2 \quad \text{whenever } x \in D \text{ and } n, m \geq N.$$

If  $m > n \geq N$  we have

$$\|F(x) - F_n(x)\| \leq \|F(x) - F_m(x)\| + \|F_m(x) - F_n(x)\| < \|F_m(x) - F(x)\| + \epsilon/2.$$

The left side of this inequality does not depend on  $m$  and the right side holds for all  $m > n$ . For each  $x \in D$ ,  $\lim_{m \rightarrow \infty} \|F(x) - F_m(x)\| = 0$ . Hence, on taking the limit of the above inequality as  $m \rightarrow \infty$ , we conclude that

$$\|F(x) - F_n(x)\| \leq \epsilon/2 < \epsilon \quad \text{for all } x \in D \text{ and } n \geq N.$$

This proves that  $\{F_n\}$  converges uniformly to  $F$  on  $D$ . □

### The Sup Norm

If  $D$  is a compact subset of  $\mathbb{R}^p$ , each continuous function  $F$  from  $D$  to  $\mathbb{R}^q$  is bounded, by Theorem 8.2.4. That is,  $\sup_D \|F(x)\|$  is finite and, in fact,  $\|F(x)\|$  actually assumes this value at some point of  $D$ . We set,

$$\|F\|_D = \sup_D \|F(x)\|.$$

This is a norm on the vector space of all continuous functions from  $D$  to  $\mathbb{R}^q$ .

**Example 8.3.8.** Find  $\|\gamma\|_I$  if  $I$  is the interval  $[0, \pi]$  and  $\gamma : I \rightarrow \mathbb{R}^2$  is the curve defined by

$$\gamma(t) = (\cos t, 1 + \sin t).$$

We have

$$\|\gamma(t)\| = \sqrt{\cos^2 t + (1 + \sin t)^2} = \sqrt{2 + 2 \sin t}.$$

This attains its maximum value on  $[0, \pi]$  at  $t = \pi/2$ , where it has the value 2. Thus,  $\|\gamma\|_I = 2$ .

**Theorem 8.3.9.** *If  $D$  is a compact subset of  $\mathbb{R}^p$  and  $\{F_n\}$  is a sequence of continuous functions from  $D$  to  $\mathbb{R}^q$ , then  $\{F_n\}$  converges uniformly to a function  $F : D \rightarrow \mathbb{R}^q$  if and only if  $\lim_{n \rightarrow \infty} \|F - F_n\|_D = 0$ .*

*Proof.* Given any  $\epsilon > 0$  and any  $n$ , the inequality  $\|F(x) - F_n(x)\| < \epsilon$  holds for all  $x \in D$  if and only if  $\|F - F_n\|_D < \epsilon$ . Thus,  $\{F_n\}$  converges uniformly to  $F$  if and only if  $\lim_{n \rightarrow \infty} \|F - F_n\|_D = 0$ .  $\square$

The space,  $\mathcal{C}(K; \mathbb{R}^q)$ , of all continuous functions on a compact set  $K \subset \mathbb{R}^p$ , with values in  $\mathbb{R}^q$  is a vector space under the operations of pointwise addition and scalar multiplication of functions. If we define the norm of an element  $F$  of this space to be the Sup norm  $\|F\|_K$ , then it is easy to see that  $\mathcal{C}(K; \mathbb{R}^q)$  is a normed vector space (Exercise 8.3.10). In particular, it is a metric space in which the distance between two elements  $F$  and  $G$  is defined to be  $\|F - G\|_K$ . It turns out that this is a complete metric space (meaning that all Cauchy sequences converge).

**Theorem 8.3.10.** *The normed vector space  $\mathcal{C}(K; \mathbb{R}^q)$  is complete.*

*Proof.* A Cauchy sequence in  $\mathcal{C}(K; \mathbb{R}^q)$ , is by definition a sequence of continuous functions which is Cauchy in the metric defined by the norm  $\|\cdot\|_K$ . Such a sequence is uniformly Cauchy on  $K$ . By Theorem 8.3.7 such a sequence converges uniformly on  $K$ . The limit function is continuous, by 8.3.4. By the previous theorem, the sequence converges in the metric defined by  $\|\cdot\|_K$  to this limit. Thus, each Cauchy sequence in the metric space  $\mathcal{C}(K; \mathbb{R}^q)$  converges to an element of  $\mathcal{C}(K; \mathbb{R}^q)$  and, hence, this space is complete.  $\square$

## Series of Functions

Given a series

$$\sum_{k=1}^{\infty} F_k(x) \tag{8.3.2}$$

whose terms  $F_k$  are functions from a domain  $D \subset \mathbb{R}^p$  into  $\mathbb{R}^q$ , we define its associated sequence of partial sums  $\{S_n\}$  in the usual way:

$$S_n(x) = \sum_{k=1}^n F_k(x).$$

The series converges pointwise if its sequence of partial sums converges pointwise, It converges uniformly on  $D$  if its sequence of partial sum converges uniformly on  $D$ .

As in the single variable case, there is a simple condition (the Weierstrass M-test) which ensures that a series converges uniformly. The proof is the same as the proof of Theorem 6.4.4 and so we will not repeat it.

**Theorem 8.3.11. (Weierstrass M-test)** *If there is a convergent series of non-negative numbers*

$$\sum_{k=1}^{\infty} M_k,$$

*such that  $\|F_k(x)\| \leq M_k$  for all  $k$  and all  $x \in D$ , then the series (8.3.2) converges uniformly on  $D$ .*

**Example 8.3.12.** Show that the series

$$\sum_{k=1}^{\infty} \frac{1}{k^2} \sin kx \cos ky \tag{8.3.3}$$

converges uniformly on  $\mathbb{R}^2$ .

**Solution:** Since

$$\left| \frac{1}{k^2} \sin kx \cos ky \right| \leq \frac{1}{k^2} \quad \text{for all } k, x, y,$$

and the series  $\sum_{k=1}^{\infty} 1/k^2$  converges (it's a p-series with  $p = 2$ ), the Weierstrass M-test tells us that the series (8.3.3) converges uniformly on  $\mathbb{R}^2$ .

### Exercise Set 8.3

1. Show that the sequence  $\{\gamma_n(t)\}$ , where

$$\gamma_n(t) = \left( \frac{1}{1+nt}, \frac{t}{n} \right)$$

does not converge uniformly on  $[0, 1]$ .

2. Show that the sequence  $\{\lambda_n(t)\}$ , where

$$\lambda_n(t) = \left( \frac{t}{1+nt}, \frac{t}{n} \right)$$

does converge uniformly on  $[0, 1]$ .

3. Does the sequence  $\{(k^{-1} \sin kx, k^{-1} \cos ky)\}$  converge pointwise on  $\mathbb{R}^2$ ? Does it converge uniformly on  $\mathbb{R}^2$ ? Justify your answers.
4. Does the sequence  $\{\sin(x/k), \cos(y/k)\}$  converge pointwise on  $\mathbb{R}^2$ ? Does it converge uniformly on  $\mathbb{R}^2$ ? Justify your answer.

5. Find  $\|F\|_D$  if  $D = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$  and  $F : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  is defined by

$$F(x, y) = (x + 1, y + 1).$$

6. Find  $\|\gamma\|_I$  if  $I = [0, \pi]$  and  $\gamma : I \rightarrow \mathbb{R}^2$  is defined by

$$\gamma(t) = (2 \cos t, 3 \sin t).$$

7. Does the series  $\sum_{k=0}^{\infty} x^k y^k$  converge uniformly on the square

$$\{(x, y) \in \mathbb{R}^2 : -1 < x < 1, -1 < y < 1\}?$$

Justify your answer.

8. Does the series  $\sum_{k=0}^{\infty} x^k y^k$  converge uniformly on the disc

$$\{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}?$$

Justify your answer.

9. Does the series  $\sum_{k=0}^{\infty} (x^k, (1-x)^k)$  converge pointwise on  $[0, 1]$ ? Does it converge pointwise on  $(0, 1)$ ? On which subsets of  $(0, 1)$  does it converge uniformly? Justify your answers.
10. If  $K$  is a compact subset of  $\mathbb{R}^p$ , show that  $\|\cdot\|_K$  is a norm on the vector space  $\mathcal{C}(K; \mathbb{R}^q)$  of continuous functions on  $K$  with values in  $\mathbb{R}^q$ .
11. Prove that if  $D$  is a subset of  $\mathbb{R}^p$  and  $\{F_n\}$  is a sequence of functions from  $D$  to  $\mathbb{R}^q$ , then  $\{F_n\}$  fails to converge uniformly to 0 if and only if there is a sequence  $\{x_n\}$  in  $D$  such that the sequence of numbers  $\{F_n(x_n)\}$  does not converge to 0.
12. If  $K \subset \mathbb{R}^p$  is compact, show that a series  $\sum_{k=1}^{\infty} F_k(x)$  of functions from  $K$  to  $\mathbb{R}^q$  converges uniformly on  $K$  if the series of numbers  $\sum_{k=1}^{\infty} \|F_k(x)\|_K$  converges.

## 8.4 Linear Functions, Matrices

Other than constants, linear functions are the simplest functions from  $\mathbb{R}^p$  to  $\mathbb{R}^q$ . For example, the linear functions from  $\mathbb{R}$  to  $\mathbb{R}$  are the functions of the form

$$L(x) = mx,$$

where  $m$  is a constant – that is, they are functions whose graphs are straight lines through the origin. In this section we introduce and study linear functions between Euclidean spaces. In the next chapter we will show how to use linear functions to approximate more complicated functions.

### Linear Functions

**Definition 8.4.1.** A function  $L : \mathbb{R}^p \rightarrow \mathbb{R}^q$  is said to be *linear* if, whenever  $x, y \in \mathbb{R}^p$  and  $a \in \mathbb{R}$ ,

- (a)  $L(x + y) = L(x) + L(y)$ ; and
- (b)  $L(ax) = aL(x)$ .

Linear functions are often called *linear transformations* or *linear operators*.

Combining (a) and (b) of this definition we see that a linear function preserves linear combinations of vectors. That is,

$$L(ax + by) = aL(x) + bL(y) \quad (8.4.1)$$

for all pairs of vectors  $x, y \in \mathbb{R}^p$  and all pairs of scalars  $a, b$ . An induction argument shows that the analogous result holds for linear combinations of more than two vectors.

Note that, since the definition uses only addition and scalar multiplication, linear functions between any two vector spaces may be defined in the same way as linear functions between  $\mathbb{R}^p$  and  $\mathbb{R}^q$ .

**Example 8.4.2.** Determine whether the functions  $F, G$  from  $\mathbb{R}^2$  to  $\mathbb{R}^2$  and  $H$  from  $\mathbb{R}^2$  to  $\mathbb{R}$  are linear, where

$$\begin{aligned} F(x, y) &= (2x + y, x - y), \\ G(x, y) &= (x^2, x + y), \\ H(x, y) &= \begin{cases} \frac{x^3 + y^3}{x^2 + y^2} & \text{if } (x, y) \neq (0, 0) \\ 0 & \text{if } (x, y) = (0, 0) \end{cases} \end{aligned}$$

**Solution:** The function  $F$  is linear since, given two vectors  $u = (x_1, y_1)$  and  $v = (x_2, y_2)$  in  $\mathbb{R}^2$  and a scalar  $a$ , we have:

$$\begin{aligned} F(u + v) &= F(x_1 + x_2, y_1 + y_2) \\ &= (2(x_1 + x_2) + (y_1 + y_2), (x_1 + x_2) - (y_1 + y_2)) \\ &= ((2x_1 + y_1) + (2x_2 + y_2), (x_1 - y_1) + (x_2 - y_2)) = F(u) + F(v) \end{aligned}$$

and

$$\begin{aligned} F(au) &= F(ax_1, ay_1) = (2(ax_1) + ay_1, ax_1 - ay_1) \\ &= (a(2x_1 + y_1), a(x_1 - y_1)) = aF(u). \end{aligned}$$

The function  $G$  is not linear since, if  $u = (1, 0)$ , then

$$G(2u) = ((2)^2, 2) = (4, 2),$$

while

$$2G(u) = 2(1^2, 1) = (2, 2).$$

These are not equal and so (b) of the above definition does not hold for  $G$ .

The function  $H$  is also not linear. If  $u = (1, 0)$  and  $v = (0, 1)$ , then

$$H(u) = H(v) = H(u + v) = 1.$$

Thus,  $H(u + v) \neq H(u) + H(v)$  and (a) of the definition does not hold (note that (b) does hold for this function).

### Linear Functions and Matrices

Recall that each vector  $x \in \mathbb{R}^p$  may be written as a linear combination of the vectors  $e_j$ , where

$$e_j = (0, \dots, 0, 1, 0, \dots, 0)$$

with the 1 in the  $j$ th place. Specifically,

$$x = \sum_{j=1}^p x_j e_j \quad (8.4.2)$$

where  $x_j$  is the  $j$ th component of the vector  $x$ .

If we apply a linear function  $L : \mathbb{R}^p \rightarrow \mathbb{R}^q$  to the vector  $x$  and use the fact that linear functions preserve linear combinations, we conclude that

$$L(x) = \sum_{k=1}^p x_k L(e_k).$$

The vector  $L(e_j) \in \mathbb{R}^q$  has  $i$ th component  $e_i \cdot L(e_j)$ . If we set

$$a_{ij} = e_i \cdot L(e_j), \quad (8.4.3)$$

then the  $i$ th component  $y_i$  of the vector  $y = L(x)$  is

$$y_i = \sum_{j=1}^p a_{ij} x_j. \quad (8.4.4)$$

The numbers  $(a_{ij})$ , appearing in (8.4.4), form a  $q \times p$  matrix – that is a rectangular array

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ a_{q1} & a_{q2} & \cdots & a_{qp} \end{pmatrix}$$

with  $q$  rows and  $p$  columns. The equation  $y = L(x)$  can be expressed in vector–matrix notation as

$$\begin{pmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ y_q \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ a_{q1} & a_{q2} & \cdots & a_{qp} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ x_p \end{pmatrix}. \quad (8.4.5)$$

In this notation, the vectors  $x$  and  $y$  are written as column vectors. The expression on the right is the vector–matrix product of the matrix  $A = (a_{ij})$  and the vector  $x = (x_j)$ . It is defined to be the vector whose  $i$ th component is the inner product of the  $i$ th row of  $A$  with the vector  $x$ .

At this point, we have shown that, to each linear function  $L : \mathbb{R}^p \rightarrow \mathbb{R}^q$ , there corresponds a  $q \times p$  matrix  $A$  such that

$$L(x) = Ax,$$

where  $Ax$  is the vector–matrix product of  $A$  with  $x$ , as in (8.4.5). On the other hand, each  $q \times p$  matrix  $A$  determines a linear function in this way, since vector matrix multiplication satisfies

$$A(x + y) = Ax + Ay \quad \text{and} \quad A(cx) = c(Ax),$$

for every pair of vectors  $x, y \in \mathbb{R}^p$  and every scalar  $c \in \mathbb{R}$  (Exercise 8.4.11).

Note that, in the correspondence between a linear function  $L$  and its matrix  $A$ , the  $j$ th column of  $A$  is the vector  $L(e_j)$ . The following theorem summarizes the above discussion.

**Theorem 8.4.3.** *A function  $L : \mathbb{R}^p \rightarrow \mathbb{R}^q$  is linear if and only if there is a  $q \times p$  matrix  $A$  such that*

$$L(x) = Ax \quad \text{for all } x \in \mathbb{R}^p.$$

**Example 8.4.4.** If a function  $L$  from  $\mathbb{R}^3$  to  $\mathbb{R}^3$  is defined by

$$L(x, y, z) = (x + 2y - z, y + z, 3x - y + z),$$

then is  $L$  linear? If so, what matrix represents it?

**Solution:** If we write  $L(x, y, z)$  as a column vector, then it clearly is given by

$$L(x, y, z) = \begin{pmatrix} x + 2y - z \\ y + z \\ 3x - y + z \end{pmatrix} = \begin{pmatrix} 1 & 2 & -1 \\ 0 & 1 & 1 \\ 3 & -1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

Since  $L$  is given by a matrix through vector–matrix multiplication, it is linear by Theorem 8.4.3.

### Matrix Operations

The sum of two linear functions  $L : \mathbb{R}^p \rightarrow \mathbb{R}^q$  and  $M : \mathbb{R}^p \rightarrow \mathbb{R}^q$  is defined pointwise, as is the sum of any two functions with a common domain. That is,  $(L + M)(x) = L(x) + M(x)$ . The function  $L + M$  is also a linear function since

$$\begin{aligned}(L + M)(x + y) &= L(x + y) + M(x + y) \\ &= L(x) + L(y) + M(x) + M(y) = (L + M)(x) + (L + M)(y),\end{aligned}$$

for all  $x, y \in \mathbb{R}^p$ , and

$$(L + M)(ax) = L(ax) + M(ax) = aL(x) + aM(x) = a(L + M)(x),$$

for all  $x \in \mathbb{R}^p$  and  $a \in \mathbb{R}$ .

Similarly, the product of a scalar  $c$  with a linear function  $L$  is defined by  $(cL)(x) = cL(x)$ . This is also, clearly, a linear function.

If  $M : \mathbb{R}^p \rightarrow \mathbb{R}^q$  and  $L : \mathbb{R}^q \rightarrow \mathbb{R}^s$  are linear functions, then the composition  $L \circ M : \mathbb{R}^p \rightarrow \mathbb{R}^s$  is defined, where

$$L \circ M(x) = L(M(x)).$$

This is also a linear function, since

$$\begin{aligned}(L \circ M)(x + y) &= L(M(x + y)) = L(M(x) + M(y)) \\ &= L(M(x)) + L(M(y)) = L \circ M(x) + L \circ M(y).\end{aligned}$$

for all  $x, y \in \mathbb{R}^q$ , and

$$L \circ M(ax) = L(M(ax)) = L(aM(x)) = aL(M(x)) = aL \circ M(x),$$

for all  $x \in \mathbb{R}^q$  and all  $a \in \mathbb{R}$ .

In view of the above, it is natural to ask, for linear functions  $L$  and  $M$  represented by matrices  $A$  and  $B$ , what are the matrices representing  $L + M$ ,  $cL$ , and  $M \circ L$ ? The answer is given in the next two theorems. They have simple proofs based on the fact that, if the matrix  $A$  represents the linear function  $L$ , then the  $j$ th row of  $A$  is  $L(e_j)$  (this is just equation (8.4.3)). The details are left to the exercises.

**Theorem 8.4.5.** *If  $L : \mathbb{R}^p \rightarrow \mathbb{R}^q$  and  $M : \mathbb{R}^p \rightarrow \mathbb{R}^q$  are linear functions represented by matrices  $A = (a_{ij})$  and  $B = (b_{ij})$ , respectively, and  $c \in \mathbb{R}$ , then  $L + M$  and  $cL$  are represented by the matrices*

$$A + B = (a_{ij} + b_{ij}) \quad \text{and} \quad cA = (ca_{ij}).$$

These are the usual operations of addition and scalar multiplication of matrices. The entry in the  $i$ th row and  $j$ th column of  $A + B$  is  $a_{ij} + b_{ij}$ , while that of  $cA$  is  $ca_{ij}$ .

**Theorem 8.4.6.** If  $L : \mathbb{R}^q \rightarrow \mathbb{R}^s$  and  $M : \mathbb{R}^p \rightarrow \mathbb{R}^q$  are linear functions represented by matrices  $A = (a_{ij})$  and  $B = (b_{jk})$ , then  $L \circ M : \mathbb{R}^p \rightarrow \mathbb{R}^s$  is represented by the matrix  $AB = (c_{ik})$ , where

$$c_{ik} = \sum_{j=1}^q a_{ij}b_{jk}.$$

This is the usual operation of matrix multiplication. The entry in the  $i$ th row and  $k$ th column of  $AB$  is the inner product of the  $i$ th row of  $A$  with the  $k$ th column of  $B$ .

**Example 8.4.7.** If  $A = \begin{pmatrix} 1 & 2 & -1 \\ 0 & 1 & 1 \end{pmatrix}$  and  $B = \begin{pmatrix} 0 & 1 & 3 \\ 1 & 0 & 1 \end{pmatrix}$ , then find  $2A - B$ .

**Solution:** We have

$$2A - B = \begin{pmatrix} 2-0 & 4-1 & -2-3 \\ 0-1 & 2-0 & 2-1 \end{pmatrix} = \begin{pmatrix} 2 & 3 & -5 \\ -1 & 2 & 1 \end{pmatrix}$$

The transpose  $A^t$  of a matrix  $A$  is the matrix obtained by interchanging the rows and columns of  $A$ . That is. If  $A = (a_{ij})$ , then  $A^t = (b_{ij})$ , where  $b_{ij} = a_{ji}$ .

**Example 8.4.8.** If  $A$  is the matrix of the previous example, then find  $A^t$ ,  $AA^t$  and  $A^tA$ .

**Solution:** By definition, we have

$$A^t = \begin{pmatrix} 1 & 0 \\ 2 & 1 \\ -1 & 1 \end{pmatrix},$$

while

$$AA^t = \begin{pmatrix} 1 & 2 & -1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 2 & 1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 6 & 1 \\ 1 & 2 \end{pmatrix}$$

and

$$A^tA = \begin{pmatrix} 1 & 0 \\ 2 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & -1 \\ 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & -1 \\ 2 & 5 & -1 \\ -1 & -1 & 2 \end{pmatrix}.$$

## Norm of a Linear Transformation

**Definition 8.4.9.** A linear transformation  $L$  from a normed vector space  $X$  to a normed vector space  $Y$  is said to be *bounded* if the set

$$\left\{ \frac{\|L(x)\|}{\|x\|} : x \in X, x \neq 0 \right\} \quad (8.4.6)$$

is bounded above. In this case, the least upper bound of this set is called the *operator norm* of  $L$  and is denoted  $\|L\|$ .

Equivalently, a linear transformation  $L$  is bounded if there is a number  $B$  such that

$$\|L(x)\| \leq B\|x\| \quad \text{for all } x \in X.$$

The operator norm  $\|L\|$  of  $L$  is the least such number  $B$ .

**Theorem 8.4.10.** *If  $X$  and  $Y$  are normed vector spaces, then every bounded linear transformation  $L : X \rightarrow Y$  is uniformly continuous on  $X$ .*

*Proof.* If  $x_1, x_2 \in X$ , then

$$\|L(x_1) - L(x_2)\| = \|L(x_1 - x_2)\| \leq \|L\| \|x_1 - x_2\|.$$

Hence, given  $\epsilon > 0$ , if we choose  $\delta = \epsilon/\|L\|$ , then

$$\|L(x_1) - L(x_2)\| \leq \|L\| \|x_1 - x_2\| < \epsilon \quad \text{whenever } \|x_1 - x_2\| < \delta.$$

This shows that  $L$  is uniformly continuous on  $X$ . □

**Theorem 8.4.11.** *Every linear transformation from  $L : \mathbb{R}^p \rightarrow \mathbb{R}^q$  is bounded and, hence, uniformly continuous. Furthermore,*

$$\|L\| \leq \left( \sum_{ij} |a_{ij}|^2 \right)^{1/2},$$

where  $A = (a_{ij})$  is the matrix which determines  $L$ .

*Proof.* Let  $A$  be the matrix which determines  $L$  and let  $r_i$  be the  $i$ th row of  $A$ . Then the  $i$ th component of  $y = L(x) = Ax$  is the inner product  $y_i = r_i \cdot x$ . By the Cauchy-Schwarz Inequality (Theorem 7.1.8)

$$|y_i| \leq \|r_i\| \|x\|.$$

Thus,

$$\begin{aligned} \|L(x)\| &= (y_1^2 + \cdots + y_q^2)^{1/2} \leq (\|r_1\|^2 + \cdots + \|r_q\|^2)^{1/2} \|x\| \\ &= \left( \sum_{ij} |a_{ij}|^2 \right)^{1/2} \|x\|. \end{aligned}$$

This implies that  $L$  is bounded and  $\|L\| \leq \left( \sum_{ij} |a_{ij}|^2 \right)^{1/2}$ . □

### Inverse of a Matrix

Of particular interest in matrix theory are *square matrices* – that is,  $p \times p$  matrices for some  $p$ . The product of two  $p \times p$  matrices is another one and so the set of  $p \times p$  matrices is closed under multiplication.

There is a multiplicative identity  $I$  in the set of  $p \times p$  matrices. This is the matrix  $I = (\delta_{ij})$  where  $\delta_{ij} = 1$  if  $i = j$  and  $\delta_{ij} = 0$  otherwise. It has the property that

$$AI = IA = A,$$

for any  $p \times p$  matrix  $A$ .

If  $A$  is a  $p \times p$  matrix, then an *inverse* for  $A$  is a  $p \times p$  matrix  $A^{-1}$  such that

$$AA^{-1} = A^{-1}A = I.$$

By Cramer's rule, a square matrix has an inverse if and only if its determinant  $\det A$  is non-zero and, in this case,

$$A^{-1} = \frac{1}{\det A}(A^c)^t,$$

where  $A^c$  is the matrix of cofactors of  $A$  – that is,  $A^c = ((-1)^{i+j} \det A_{ij})$ , where  $A_{ij}$  is the  $(p-1) \times (p-1)$  matrix obtained by deleting the  $i$ th row and  $j$ th column from  $A$ .

A matrix is said to be *non-singular* if it has an inverse, that is, if its determinant is non-zero. A square matrix is *singular* if it fails to have an inverse.

Note that if  $L : \mathbb{R}^d \rightarrow \mathbb{R}^d$  is a linear transformation with matrix  $A$ , then  $A$  has an inverse matrix  $A^{-1}$  if and only if  $L$  has an inverse transformation  $L^{-1}$  and, in this case, the linear transformation  $L^{-1}$  has  $A^{-1}$  as its associated matrix.

**Example 8.4.12.** Let

$$A = \begin{pmatrix} 2 & 1 \\ -1 & 1 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 2 & -1 \\ -2 & 1 \end{pmatrix}.$$

For each of  $A$  and  $B$ , determine if the matrix has an inverse and, if it does, find it.

**Solution:** The matrices  $A$  and  $B$  have determinants

$$\det A = 2 + 1 = 3 \quad \text{and} \quad \det B = 2 - 2 = 0.$$

Thus,  $A$  has an inverse and  $B$  does not. By Cramer's rule, the inverse of  $A$  is

$$\frac{1}{3} \begin{pmatrix} 1 & -1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 1/3 & -1/3 \\ 1/3 & 2/3 \end{pmatrix}.$$

**Remark 8.4.13.** In what follows, we will often ignore the difference between a linear function  $L$  and the matrix which represents it. They are not exactly the same. The matrix of a linear transformation depends on a choice of coordinate systems in  $\mathbb{R}^p$  and  $\mathbb{R}^q$ , while the linear transformation is independent of the

choice of coordinates. To ignore the distinction will not cause problems as long as we stick with one coordinate system. There will, however, be occasions where we change coordinate systems in  $\mathbb{R}^p$  or  $\mathbb{R}^q$  or both while dealing with a given linear transformation. It should be understood that the matrix corresponding to the linear transformation will, as a result, also change.

### Exercise Set 8.4

The first five exercises involve the matrices

$$A = \begin{pmatrix} 3 & -1 \\ 2 & 1 \end{pmatrix}, B = \begin{pmatrix} 2 & 5 \\ -2 & 2 \end{pmatrix}, C = \begin{pmatrix} 1 & -1 \\ 4 & -6 \\ -1 & 2 \end{pmatrix}, D = \begin{pmatrix} 2 & 0 & 1 \\ -1 & 1 & 3 \end{pmatrix}.$$

1. Find  $2A + B$ ,  $A - B$ ,  $AB$  and  $BA$ .
2. Find  $\det A$  and  $\det B$  and  $A^{-1}$  and  $B^{-1}$ .
3. Find  $CD$  and  $DC$ .
4. Based on the result of the previous exercise, can you tell what  $(CD)^2$  is without doing any further calculation?
5. Find  $\det CD$ .
6. Is the function  $F : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  defined by  $F(x, y) = (x + y, xy)$  a linear transformation? If so, what is its matrix?
7. Is the the function  $F : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  defined by  $F(x, y) = (x + y, x - y)$  a linear transformation? If so, what is its matrix?
8. Is the transformation of  $\mathbb{R}^2$  to itself which rotates every vector through an angle  $\theta$  (counterclockwise rotations have positive angle and clockwise rotations have negative angle) a linear transformation? If so, what is its matrix?
9. What is the matrix for the linear transformation of  $\mathbb{R}^2$  which reflects each point through the diagonal line  $y = x$  (this transformation interchanges the  $x$  and  $y$  coordinates of each point).
10. Find a linear transformation  $L : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  such that  $L(1, 2, 1) = (1, 2, 1)$  and  $L(u) = 0$  for every vector  $u \in \mathbb{R}^3$  which is orthogonal to  $(1, 2, 1)$ .
11. Prove that if  $A$  is a  $q \times p$  matrix, then

$$A(x + y) = Ax + Ay \quad \text{and} \quad A(cx) = c(Ax),$$

for every pair of vectors  $x, y \in \mathbb{R}^p$  and every scalar  $c \in \mathbb{R}$ .

12. Prove Theorem 8.4.5.

13. Prove Theorem 8.4.6.
14. Prove that if  $K$  and  $L$  are linear transformations from  $\mathbb{R}^p \rightarrow \mathbb{R}^q$ , then

$$\|K + L\| \leq \|K\| + \|L\|.$$

15. Prove that if  $K : \mathbb{R}^p \rightarrow \mathbb{R}^q$  and  $L : \mathbb{R}^q \rightarrow \mathbb{R}^r$  are linear transformations, then

$$\|L \circ K\| \leq \|L\| \|K\|.$$

16. Prove that the operator norm of a  $p \times p$  diagonal matrix has norm equal to the largest absolute value of the elements on the diagonal.

## 8.5 Dimension, Rank, Lines, and Planes

A vector space  $X$  has finite dimension if it contains a finite set  $\{x_1, x_2, \dots, x_k\}$  of vectors which span  $X$  – that is, every vector in  $X$  is a linear combination of the vectors  $x_j$ . If this set is also *linearly independent*, meaning the only linear combination of the vectors  $x_j$  that equals 0 is the one in which all coefficients are zero, then the set  $\{x_1, x_2, \dots, x_k\}$  is called a *basis* for  $X$ . In this case, each element of  $X$  is a unique linear combination of the vectors  $x_j$ . Every finite dimensional vector space  $X$  has a basis. In fact  $X$  has many bases, but each of them has the same number of elements. This number is called the *dimension* of  $X$  and written  $\dim(X)$ .

A subset  $M$  of a vector space  $X$  is called a *linear subspace* if it is closed under addition and scalar multiplication – that is,  $x + y \in M$  and  $ax \in M$  whenever  $x, y \in M$  and  $a \in \mathbb{R}$ . It follows that a linear subspace  $M$  of a vector space is itself a vector space, with addition and scalar multiplication in  $M$  defined in the same way they are defined in  $X$ . If  $X$  is finite dimensional, then so is the subspace  $M$  and any basis  $\{x_1, x_2, \dots, x_m\}$  for  $M$  can be expanded to a basis  $\{x_1, x_2, \dots, x_m, x_{m+1}, \dots, x_n\}$  for  $X$ . Thus

$$\dim(M) \leq \dim(X).$$

The set  $\{e_1, \dots, e_p\}$  is a basis for  $\mathbb{R}^p$ , where recall that  $e_j$  is the  $p$ -tuple which has 1 for its  $j$ th component and 0 for all the others. However, this is not the only basis for  $\mathbb{R}^p$ .

**Example 8.5.1.** Show that the vectors  $u = (1, 0, 1)$ ,  $v = (1, 1, 0)$ , and  $w = (0, 1, 1)$  form a basis for  $\mathbb{R}^3$ .

**Solution:** Consider the vector equation

$$au + bv + cw = y. \tag{8.5.1}$$

To show that  $\{u, v, w\}$  spans  $\mathbb{R}^3$ , we must show that this equation has a solution for every  $y$ . To show that  $\{u, v, w\}$  is a linearly independent set, we must show that if  $y = 0$ , then this equation has only the zero solution for  $(a, b, c)$ . Taken

together, these two statements mean that equation (8.5.1) should have a unique solution for every  $y \in \mathbb{R}^3$ . The vector equation (8.5.1) is equivalent to the system of linear equations

$$\begin{aligned} a + b + 0 &= y_1 \\ 0 + b + c &= y_2, \\ a + 0 + c &= y_3 \end{aligned}$$

which, in turn, may be written as the vector matrix equation

$$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}.$$

The matrix in this equation has determinant 2 and so the matrix has an inverse. This implies that the equation has a unique solution  $(a, b, c)$  for each  $y = (y_1, y_2, y_3)$  and, hence, that  $\{u, v, w\}$  is a basis for  $\mathbb{R}^3$ .

**Definition 8.5.2.** If  $L : X \rightarrow Y$  is a linear transformation between vector spaces, then the *image* of  $L$ , denoted  $\text{im}(L)$  is the set

$$L(X) = \{L(x) : x \in X\},$$

while the *kernel* of  $L$ , denoted  $\ker(L)$ , is the set

$$\{x \in X : L(x) = 0\}.$$

Since  $L$  is linear, it follows easily that its kernel and image are linear subspaces of  $X$  and  $Y$ , respectively.

**Theorem 8.5.3.** *If  $L : X \rightarrow Y$  is a linear transformation between finite dimensional vector spaces, then*

$$\dim(\ker(L)) + \dim(\text{im}(L)) = \dim(X).$$

*Proof.* Let  $\dim(\ker(L)) = m$  and let  $\{x_1, x_2, \dots, x_m\}$  be a basis for  $\ker(L)$ . We may expand this to a basis  $\{x_1, x_2, \dots, x_m, x_{m+1}, \dots, x_n\}$  for  $X$ .

Set  $y_j = L(x_{m+j})$  for  $j = 1, \dots, n - m$ . Since every vector in  $X$  is a linear combination of the vectors  $x_1, \dots, x_n$  and  $L(x_k) = 0$  for  $k = 1, \dots, m$ , we conclude that every vector in  $\text{im}(L)$  is a linear combination of the vectors  $y_1, \dots, y_{n-m}$ . This set of vectors is linearly independent, since if

$$a_1 y_1 + a_2 y_2 + \dots + a_{n-m} y_{n-m} = 0,$$

then  $a_1 x_{m+1} + a_2 x_{m+2} + \dots + a_{n-m} x_n \in \ker(L)$ . This implies that there are numbers  $b_1, \dots, b_m$  such that

$$a_1 x_{m+1} + a_2 x_{m+2} + \dots + a_{n-m} x_n = b_1 x_1 + b_2 x_2 + \dots + b_m x_m.$$

However, since  $\{x_1, \dots, x_n\}$  is a linearly independent set, the  $a_j$ s and  $b_k$ s must all be 0. The fact that the  $a_j$ s must all be 0 shows that the set  $\{y_1, \dots, y_{n-m}\}$  is linearly independent and, hence, forms a basis for  $\text{im}(L)$ .

We now have  $\dim(X) = n$ ,  $\dim(\ker(L)) = m$  and  $\dim(\text{im}(L)) = n - m$ . Thus,  $\dim(\ker(L)) + \dim(\text{im}(L)) = \dim(X)$ , as claimed.  $\square$

**Definition 8.5.4.** Let  $A$  be a  $q \times p$  matrix and let  $L : \mathbb{R}^p \rightarrow \mathbb{R}^q$  be the linear transformation it determines. Then  $\text{Rank}(A)$  is defined to be  $\dim(\text{im}(L))$ . Equivalently, by the previous theorem, it is also equal to  $\dim(X) - \dim(\ker(L))$ . If  $L$  is a linear transformation whose matrix has rank  $r$ , then we will also say that  $L$  has rank  $r$ .

A *submatrix* of a matrix  $A$  is a matrix obtained from  $A$  by deleting some of its rows and columns.

The following is proved in most linear algebra texts. We won't repeat the proof here.

**Theorem 8.5.5.** *The rank of a  $q \times p$  matrix  $A$  is  $r$ , where  $r \times r$  is the dimension of the largest square submatrix of  $A$  with non-zero determinant.*

**Example 8.5.6.** What is the rank of the matrix

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \\ 1 & -1 \end{pmatrix}?$$

**Solution:** This matrix has

$$\begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix}$$

as a  $2 \times 2$  submatrix with determinant  $-3$ . It has no square submatrices of larger dimension. Therefore, the matrix  $A$  has rank 2.

**Example 8.5.7.** What is the rank of the matrix

$$B = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & -1 & -2 \end{pmatrix}?$$

**Solution:** This matrix also has

$$\begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix}$$

as a  $2 \times 2$  submatrix with determinant  $-3$ . The only square submatrix of larger dimension is the matrix  $B$  and this has determinant 0. Therefore, the matrix  $B$  also has rank 2.

## Affine Functions

**Definition 8.5.8.** An *affine function*  $F : \mathbb{R}^p \rightarrow \mathbb{R}^q$  is a function of the form

$$F(x) = b + L(x),$$

where  $b \in \mathbb{R}^q$  and  $L : \mathbb{R}^p \rightarrow \mathbb{R}^q$  is a linear function. The *rank* of an affine transformation  $F$  is the rank of its linear part  $L$ .

An *affine subspace*  $M$  of  $\mathbb{R}^p$  is a translate  $b + N$  of a linear subspace  $N$  of  $\mathbb{R}^p$ . In this case, the dimension of  $M$  is defined to be the dimension of  $N$ .

The image of an affine function  $F(x) = b + L(x)$  is an affine subspace  $b + \text{im}(L)$  – that is, it is the translate  $b + \text{im}(L)$  of the linear subspace  $\text{im}(L)$ . The dimension of this subspace is the rank of  $L$ .

Similarly, if  $F(x) = b + L(x)$  is an affine function, then the set of solutions to the vector equation  $F(x) = 0$  is also an affine subspace. In fact, if  $a$  is one such solution (so that  $F(a) = b + L(a) = 0$ ), then  $x$  is also a solution if and only if

$$L(x - a) = -b + b = 0.$$

Hence,  $x$  is a solution if and only if  $x \in a + \ker(L)$ . Thus, the set of solutions of the vector equation  $F(x) = 0$  is the translate  $a + \ker(L)$  of the linear subspace  $\ker(L)$  of  $\mathbb{R}^p$  and, hence, is an affine subspace. The dimension of this subspace is  $p - \text{Rank}(L)$ .

In general, if  $M = b + N$  an affine subspace, with  $N$  the corresponding linear subspace, then we will say that  $N$  is the set of vectors *parallel* to the affine subspace  $M$ .

### Lines in $\mathbb{R}^3$

Lines in  $\mathbb{R}^p$  are one dimensional affine subspaces of  $\mathbb{R}^p$ . The above discussion suggests expressing them as either images of rank 1 affine transformations or as kernels of rank  $p - 1$  affine transformations with domain  $\mathbb{R}^p$ .

A rank 1 affine transformation  $\gamma : \mathbb{R} \rightarrow \mathbb{R}^q$  has the form

$$\gamma(t) = a + tu. \tag{8.5.2}$$

The image of this transformation is a line which contains the point  $a = \gamma(0)$  and is parallel to the vector  $u = \gamma(1) - \gamma(0)$ .

On the other hand, given a line in  $\mathbb{R}^q$ , if we choose distinct points  $a$  and  $b$  on the line, and we set  $u = b - a$ , then the image of the affine transformation (8.5.2) is a line which contains both  $a = \gamma(0)$  and  $b = \gamma(1)$  and, hence, is the line we started with.

Thus, the lines in  $\mathbb{R}^q$  are exactly the images of affine transformations of the form (8.5.2). This situation is often expressed as a vector equation

$$x = a + tu,$$

which describes the points  $x$  on the line as the values assumed by the right side of the equation as  $t$  ranges over  $\mathbb{R}$ . This is a *parametric vector equation* for the line.

In  $\mathbb{R}^3$ , a parametric vector equation for a line takes the form  $(x, y, z) = (a_1, a_2, a_3) + t(u_1, u_2, u_3)$ , which is equivalent to the system of parametric equations

$$\begin{aligned} x &= a_1 + tu_1 \\ y &= a_2 + tu_2 \\ z &= a_3 + tu_3 \end{aligned}$$

**Example 8.5.9.** Find parametric equations for the line in  $\mathbb{R}^3$  which contains the point  $(1, 0, 0)$  and is parallel to the vector  $u = (-3, 4, 5)$ .

**Solution:** A parametric vector equation for this line is

$$(x, y, z) = (1, 0, 0) + t(-3, 4, 5).$$

The corresponding system of parametric equations is

$$\begin{aligned} x &= 1 - 3t \\ y &= 4t \\ z &= 5t \end{aligned}$$

**Example 8.5.10.** Find parametric equations for the line in  $\mathbb{R}^3$  containing the points  $(2, 1, 1)$  and  $(5, -1, 3)$ .

**Solution:** If we set  $u = (5, -1, 3) - (2, 1, 1) = (3, -2, 2)$ , then the parametric equation for our line in vector form is

$$(x, y, z) = (2, 1, 1) + t(3, -2, 2) = (2 + 3t, 1 - 2t, 1 + 2t).$$

This can also be expressed as the system of parametric equations

$$\begin{aligned} x &= 2 + 3t \\ y &= 1 - 2t \\ z &= 1 + 2t \end{aligned}$$

To express a line in  $\mathbb{R}^q$  as the kernel of an affine transformation, we choose a point  $a$  on the line and a vector  $u$  parallel to the line (we may choose  $u = b - a$  where  $b$  is a point on the line distinct from  $a$ ). If  $A$  is a matrix whose rows form a basis for the linear subspace

$$\{y \in \mathbb{R}^p : y \cdot u = 0\},$$

then  $A$  is a  $p-1 \times p$  matrix of rank  $p-1$  and  $Au = 0$ . This means that the kernel of the linear transformation determined by  $A$  has dimension 1 and contains  $u$ . Hence, this kernel is  $\{tu : t \in \mathbb{R}\}$ . The line  $\{a + tu : t \in \mathbb{R}\}$  contains  $a$  and is parallel to  $u$ . Thus, it must be our original line. By the construction of  $A$ , it also has the form

$$\{x \in \mathbb{R}^p : A(x - a) = 0\} = \{x \in \mathbb{R}^p : Ax - c = 0\} \quad \text{where } c = Aa.$$

Thus, our line is the kernel of the affine transformation  $F$  defined by  $F(x) = Ax - c$ .

If we apply the above discussion to  $\mathbb{R}^3$ , we conclude that the typical line in  $\mathbb{R}^3$  is the set of solutions  $(x, y, z)$  to an equation of the form

$$\begin{pmatrix} v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix},$$

Where  $(v_1, v_2, v_3)$  and  $(w_1, w_2, w_3)$  are linearly independent vectors. In other words, it is the set of all simultaneous solutions of the pair of linear equations

$$\begin{aligned}v_1x + v_2y + v_3z &= c_1 \\w_1x + w_2y + w_3z &= c_2.\end{aligned}$$

**Example 8.5.11.** Express the line in Example 8.5.10 as the set of solutions of a pair of linear equations.

**Solution:** We need to find two linearly independent vectors which are orthogonal to  $u = (3, -2, 2)$ . Such a pair is  $(2, 3, 0)$  and  $(2, 1, -2)$ . If we apply the matrix with these two vectors as rows to the vector  $a = (2, 1, 1)$ , the result is

$$\begin{pmatrix} 2 & 3 & 0 \\ 2 & 1 & -2 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 7 \\ 3 \end{pmatrix},$$

Thus, in vector matrix form, the equation of our line is

$$\begin{pmatrix} 2 & 3 & 0 \\ 3 & 1 & -2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 7 \\ 3 \end{pmatrix},$$

This is equivalent to the pair of simultaneous equations

$$\begin{aligned}2x + 3y &= 7 \\3x + y - 2z &= 3\end{aligned}$$

### Planes in $\mathbb{R}^3$

A plane in  $\mathbb{R}^p$  is a two dimensional affine subspace of  $\mathbb{R}^p$  – that is, a translate of a two dimensional linear subspace of  $\mathbb{R}^p$ . Such an object can be described as the image of an affine transformation of rank 2 or the kernel of an affine transformation of rank  $p - 2$  with domain  $\mathbb{R}^p$ .

If  $u$  and  $v$  are linearly independent vectors in  $\mathbb{R}^p$ , then they form a basis for a 2-dimensional linear subspace of  $\mathbb{R}^p$ . If we translate this subspace by adding  $a$  to each of its points, we obtain a plane which contains  $a$  and is parallel to  $u$  and  $v$ . It consists of all points of the form

$$x = a + su + tv;$$

that is, it is the image of the affine transformation  $F : \mathbb{R}^2 \rightarrow \mathbb{R}^p$  defined by

$$F(s, t) = a + su + tv.$$

This is the vector parametric form for the equation of a plane.

In the case where  $p = 3$ , a vector parametric equation of a plane has the form

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} + \begin{pmatrix} u_1 & v_1 \\ u_2 & v_2 \\ u_3 & v_3 \end{pmatrix} \begin{pmatrix} s \\ t \end{pmatrix},$$

or, when written as a system of equations,

$$\begin{aligned}x &= a_1 + su_1 + tv_1 \\y &= a_2 + su_2 + tv_2 \\z &= a_3 + su_3 + tv_3\end{aligned}$$

Given three points  $a, b, c$  in  $\mathbb{R}^p$  which do not lie on the same line, the vectors  $u = b - a$  and  $v = c - a$  are linearly independent (Exercise 8.5.15). Hence,  $a, u,$  and  $v$  determine an affine function  $F$  with image a plane, as above. This plane contains the points  $a = F(0, 0), b = F(1, 0),$  and  $c = F(0, 1).$

**Example 8.5.12.** Find parametric equations for the plane that contains the three points  $(1, 0, 1), (1, 1, 2), (-1, 2, 0).$

**Solution:** We choose  $a = (1, 0, 1), u = (1, 1, 2) - (1, 0, 1) = (0, 1, 1),$  and  $v = (-1, 2, 0) - (1, 0, 1) = (-2, 2, -1).$  Then, according to the above discussion, the plane we seek has parametric equations

$$\begin{aligned}x &= 1 - 2t \\y &= s + 2t \\z &= 1 + s - t.\end{aligned}$$

We can also express a plane in  $\mathbb{R}^3$  as the kernel of a rank 1 affine transformation from  $\mathbb{R}^3$  to  $\mathbb{R}$ . If  $a = (a_1, a_2, a_3)$  is a fixed point in the plane,  $u = (x, y, z)$  the general point of the plane, and  $v = (v_1, v_2, v_3)$  a vector perpendicular to the plane, then  $v \cdot (u - a) = 0$ . Thus, the plane is the kernel of the affine transformation  $f : \mathbb{R}^3 \rightarrow \mathbb{R}$  defined by  $f(u) = v \cdot u - b,$  where  $b = v \cdot a.$  The equation of the plane is then

$$v_1x + v_2y + v_3z = b.$$

**Example 8.5.13.** Find an equation for the plane of Example 8.5.12.

**Solution:** We choose  $a = (1, 0, 1)$  as a point in the plane. Now we need a vector perpendicular to the plane. The vectors  $(0, 1, 1)$  and  $(-2, 2, -1)$  are parallel to the plane and so we need to find a vector orthogonal to each of these. In fact,  $(3, 2, -2)$  is orthogonal to each of these vectors. Also,

$$(3, 2, -2) \cdot (1, 0, 1) = 1.$$

Hence, an equation for our plane is

$$3x + 2y - 2z = 1.$$

### Exercise Set 8.5

1. Do the vectors  $(1, 2, 1), (2, 0, 1),$  and  $(1, -1, 1)$  form a basis for  $\mathbb{R}^3$ . Justify your answer.
2. Do the vectors  $(1, 2, 1), (2, 0, 1),$  and  $(0, 4, 1)$  form a basis for  $\mathbb{R}^3$ . Justify your answer.

3. What is the rank of the matrix  $\begin{pmatrix} 1 & 2 & 1 \\ 2 & 3 & -1 \\ 1 & 1 & -2 \end{pmatrix}$ ?
4. What is the rank of the matrix  $\begin{pmatrix} 1 & -2 & 3 \\ -2 & 4 & -6 \end{pmatrix}$ ?
5. What is the rank of the matrix  $\begin{pmatrix} 1 & -2 & 3 \\ -2 & 3 & -6 \end{pmatrix}$ ?
6. Find parametric equations for the line in  $\mathbb{R}^3$  which contains the point  $(1, 2, 3)$  and is parallel to the vector  $(1, 1, 1)$ .
7. Find parametric equations for the line in  $\mathbb{R}^3$  containing both  $(1, 1, 1)$  and  $(3, -1, 3)$ .
8. Express the line of the previous exercise as the set of simultaneous solutions of a pair of linear equations.
9. Find parametric equations for the plane that contains the three points  $(1, 0, -1)$ ,  $(2, 1, 2)$ ,  $(-1, 2, 3)$ .
10. Express the plane of the previous exercise as the set of solutions of a linear equation.
11. Find parametric equations for a line which passes through the origin and is perpendicular to the plane  $x - y + 3z = 5$ . Use this line to determine the distance from the plane to the origin.
12. Find the distance from the line with parametric vector equation  $(x, y, z) = (1 + 2t, 2 - t, 4 + t)$  to the origin.
13. Find a formula for the point on the one dimensional subspace of  $\mathbb{R}^p$  generated by a non-zero vector  $u$  which is closest to the point  $a \in \mathbb{R}^p$ .
14. Prove that, in  $\mathbb{R}^3$ , a plane and a line not parallel to it must meet in exactly one point.
15. Prove that if  $a$ ,  $b$ , and  $c$  are three points in  $\mathbb{R}^p$  which do not lie on the same line, then the vectors  $u = b - a$  and  $v = c - a$  are linearly independent.
16. Prove that if  $M$  is a linear subspace of  $\mathbb{R}^p$  and we set

$$M^\perp = \{y \in \mathbb{R}^p : y \perp x \text{ for all } x \in M\},$$

then  $M^\perp$  is also a linear subspace of  $\mathbb{R}^p$  and every vector in  $u \in \mathbb{R}^p$  may be written in a unique way as  $u = x + y$  with  $x \in M$  and  $y \in M^\perp$  (see Definition 7.1.9).