CHAPTER 8

Euclidean Spaces

The world we live in is at least four dimensional: three spatial dimensions together with the time dimension. Moreover, certain problems from engineering, physics, chemistry, and economics force us to consider even higher dimensions. For example, guidance systems for missiles frequently require as many as 100 variables (longitude, latitude, altitude, velocity, time after launch, pitch, yaw, fuel on board, etc.). Another example, the state of a gas in a closed container, can best be described by a function of \(6n\) variables, where \(n\) is the number of molecules in the system. (Six enters the picture because each molecule of gas is described by three space variables and three momentum variables.) Thus, there are practical reasons for studying functions of more than one variable.

8.1 ALGEBRAIC STRUCTURE

For each \(n \in \mathbb{N}\), let \(\mathbb{R}^n\) denote the \(n\)-fold cartesian product of \(\mathbb{R}\) with itself; that is,

\[
\mathbb{R}^n := \{(x_1, x_2, \ldots, x_n) : x_j \in \mathbb{R} \text{ for } j = 1, 2, \ldots, n\}.
\]

By a Euclidean space we shall mean \(\mathbb{R}^n\) together with the "Euclidean inner product" defined in Definition 8.1 below. The integer \(n\) is called the dimension of \(\mathbb{R}^n\), elements \(x = (x_1, x_2, \ldots, x_n)\) of \(\mathbb{R}^n\) are called points or vectors or ordered \(n\)-tuples, and the numbers \(x_j\) are called coordinates, or components, of \(x\). Two vectors \(x, y\) are said to be equal if and only if their components are equal; that is, if and only if \(x_j = y_j\) for \(j = 1, 2, \ldots, n\). The zero vector is the vector whose components are all zero; that is, \(0 := (0, 0, \ldots, 0)\). When \(n = 2\) (respectively, \(n = 3\)), we usually denote the components of \(x\) by \(x\) and \(y\) (respectively, by \(x, y, z\)).

You have already encountered the sets \(\mathbb{R}^n\) for small \(n\). \(\mathbb{R}^1 = \mathbb{R}\) is the real line; we shall call its elements scalars. \(\mathbb{R}^2\) is the \(xy\)-plane used to graph functions of the form \(y = f(x)\). And \(\mathbb{R}^3\) is the \(xyz\)-space used to graph functions of the form \(z = f(x, y)\).

We have called elements of \(\mathbb{R}^n\) points and vectors. In general, we make no distinction between points and vectors, but in each situation we adopt the interpretation which proves most useful.

In earlier courses, vectors were (most likely) directed line segments, but our vectors look like points in \(\mathbb{R}^n\). What is going on? When we call an \(a \in \mathbb{R}^n\) a vector, we are thinking of the directed-line-segment which starts at the origin and ends at the point \(a\).
What about directed line segments which begin at arbitrary points? Two arbitrary directed line segments are said to be equivalent if and only if they have the same length and same direction. Thus every directed line segment \( V \) is equivalent to a directed line segment in standard position; that is, one which points in the same direction as \( V \), has the same length as \( V \), but whose "tail" sits at the origin and whose "head," \( \mathbf{a} \), is a point in \( \mathbb{R}^n \). If we identify \( V \) with \( \mathbf{a} \), then we can represent any arbitrary directed line segment in \( \mathbb{R}^n \) by a point in \( \mathbb{R}^n \).

Identifying arbitrary vectors in \( \mathbb{R}^n \) with vectors in standard position and, in turn, with points in \( \mathbb{R}^n \) may sound confusing and sloppy, but it is no different from letting \( 1/2 \) represent \( 2/4, 3/6, 4/8 \), and so on. (In both cases, there is an underlying equivalence relation, and we are using one member of an equivalence class to represent all of its members. For vectors, we are using the representative which lies in standard position; for rational numbers, we are using the representative which is in reduced form.)

We began our study of functions of one variable by examining the algebraic structure of \( \mathbb{R} \). In this section we begin our study of functions of several variables by examining the algebraic structure of \( \mathbb{R}^n \). That structure is described in the following definition.

**8.1 Definition.**

Let \( \mathbf{x} = (x_1, \ldots, x_n) \), \( \mathbf{y} = (y_1, \ldots, y_n) \) \( \in \mathbb{R}^n \), and \( \alpha \in \mathbb{R} \).

i) The **sum** of the vectors \( \mathbf{x} \) and \( \mathbf{y} \) is the vector

\[
\mathbf{x} + \mathbf{y} := (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n).
\]

ii) The **difference** of the vectors \( \mathbf{x} \) and \( \mathbf{y} \) is the vector

\[
\mathbf{x} - \mathbf{y} := (x_1 - y_1, x_2 - y_2, \ldots, x_n - y_n).
\]

iii) The **product** of the scalar \( \alpha \) and the vector \( \mathbf{x} \) is the vector

\[
\alpha \mathbf{x} := (\alpha x_1, \alpha x_2, \ldots, \alpha x_n).
\]

iv) The **(Euclidean) dot product** (or **scalar product** or **inner product**) of the vectors \( \mathbf{x} \) and \( \mathbf{y} \) is the scalar

\[
\mathbf{x} \cdot \mathbf{y} := x_1y_1 + x_2y_2 + \cdots + x_ny_n.
\]

These algebraic operations are analogues of addition, subtraction, and multiplication on \( \mathbb{R} \). It is natural to ask, Do the usual laws of algebra hold in \( \mathbb{R}^n \)? An answer to this question is contained in the following result.
8.2 Theorem. Let $x, y, z \in \mathbb{R}^n$ and $\alpha, \beta \in \mathbb{R}$. Then

$$
\begin{align*}
\alpha 0 &= 0, \quad 0x = 0, \quad 0 \cdot x = 0, \quad 1x = x, \quad 0 + x = x, \quad x - x = 0, \\
\alpha(\beta x) &= \beta(\alpha x) = (\alpha\beta)x, \quad \alpha(x \cdot y) = (\alpha x) \cdot y = x \cdot (\alpha y), \\
x + y &= y + x, \quad x + (y + z) = (x + y) + z, \quad x \cdot y = y \cdot x, \\
\alpha(x + y) &= \alpha x + \alpha y, \quad \text{and} \quad x \cdot (y + z) = x \cdot y + x \cdot z.
\end{align*}
$$

Proof. These properties are direct consequences of Definition 8.1 and corresponding properties of real numbers. We will prove that vector addition is associative, and leave the proof of the rest of these properties as an exercise.

By definition and associativity of addition on $\mathbb{R}$ (see Postulate 1 in Section 1.2),

$$
x + (y + z) = (x_1, \ldots, x_n) + (y_1 + z_1, \ldots, y_n + z_n)
= (x_1 + (y_1 + z_1), \ldots, x_n + (y_n + z_n))
= ((x_1 + y_1) + z_1, \ldots, (x_n + y_n) + z_n) = (x + y) + z.
$$

Thus (with the exception of the closure of the dot product and the existence of the multiplicative identity and multiplicative inverses), $\mathbb{R}^n$ satisfies the same algebraic laws, listed in Postulate 1, that $\mathbb{R}$ does. This means one can use instincts developed in high school algebra to compute with these vector operations. For example, just as $(x - y)^2 = x^2 - 2xy + y^2$ holds for real numbers $x$ and $y$, even so,

$$
(x - y) \cdot (x - y) = x \cdot x - 2x \cdot y + y \cdot y
$$

holds for any vectors $x, y \in \mathbb{R}^n$.

In the first four chapters, we used algebra together with the absolute value to define convergence of sequences and functions in $\mathbb{R}$. Is there an analogue of the absolute value for $\mathbb{R}^n$? The following definition illustrates the fact that there are many such analogues.

8.3 Definition.

Let $x \in \mathbb{R}^n$.

i) The (Euclidean) norm (or magnitude) of $x$ is the scalar

$$
\|x\| := \sqrt{\sum_{k=1}^{n} |x_k|^2}.
$$

ii) The $\ell_1$-norm (read L-one-norm) of $x$ is the scalar

$$
\|x\|_1 := \sum_{k=1}^{n} |x_k|.
$$
8.3 Definition. (Continued)

iii) The sup-norm of \( \mathbf{x} \) is the scalar

\[
\|\mathbf{x}\|_\infty := \max\{|x_1|, \ldots, |x_n|\}.
\]

iv) The (Euclidean) distance between two points \( \mathbf{a}, \mathbf{b} \in \mathbb{R}^n \) is the scalar

\[
\text{dist}(\mathbf{a}, \mathbf{b}) := \|\mathbf{a} - \mathbf{b}\|.
\]

(Note: For relationships between these three norms, see Remark 8.7 below. The subscript \( \infty \) is frequently used for supremum norms because the supremum of a continuous function on an interval \([a, b]\) can be computed by taking the limit of \( \left( \int_a^b |f(x)|^p \, dx \right)^{1/p} \) as \( p \to \infty \)—see Exercise 5.2.8.)

Since \( \|\mathbf{x}\| = \|\mathbf{x}\|_1 = \|\mathbf{x}\|_\infty = |x| \), when \( n = 1 \), each norm defined above is an extension of the absolute value from \( \mathbb{R} \) to \( \mathbb{R}^n \). The most important, and in some senses the most natural, of these norms is the Euclidean norm. This is true for at least two reasons. First, by definition,

\[
\|\mathbf{x}\|^2 = \mathbf{x} \cdot \mathbf{x} \quad \text{for all} \quad \mathbf{x} \in \mathbb{R}^n.
\]

(This aids in many calculations; see, for example, the proofs of Theorems 8.5 and 8.6 below.) Second, if \( \Delta \) is the triangle in \( \mathbb{R}^2 \) with vertices \((0, 0), (a, b), \) and \((a, 0)\), then by the Pythagorean Theorem, the hypotenuse of \( \Delta \), \( \sqrt{a^2 + b^2} \), is exactly the norm of \( \mathbf{x} \). In particular, the Euclidean norm of a vector has a simple geometric interpretation in \( \mathbb{R}^2 \).

The algebraic structure of \( \mathbb{R}^n \) also has a simple geometric interpretation in \( \mathbb{R}^2 \) which gives us another very useful way to think about vectors. To describe it, fix vectors \( \mathbf{a} = (a_1, a_2) \) and \( \mathbf{b} = (b_1, b_2) \) and let \( \mathcal{P}(\mathbf{a}, \mathbf{b}) \) denote parallelogram associated with \( \mathbf{a} \) and \( \mathbf{b} \) (i.e., the parallelogram whose sides are given by \( \mathbf{a} \) and \( \mathbf{b} \)). (We are assuming that this parallelogram is not degenerate—see Figure 8.1.)

Then the vector sum of \( \mathbf{a} \) and \( \mathbf{b} \), \((a_1 + b_1, a_2 + b_2)\), is evidently the diagonal of \( \mathcal{P}(\mathbf{a}, \mathbf{b}) \); that is, \( \mathbf{a} + \mathbf{b} \) is the vector which begins at the origin and ends at the opposite vertex of \( \mathcal{P}(\mathbf{a}, \mathbf{b}) \). Similarly, the difference \( \mathbf{a} - \mathbf{b} \) can be identified with the other diagonal of \( \mathcal{P}(\mathbf{a}, \mathbf{b}) \) (see Figure 8.1). The scalar product of \( t \) and \( \mathbf{a} = (a_1, a_2) \), evidently stretches or compresses the vector \( \mathbf{a} \), but leaves it in the same straight line which passes through \( 0 \) and \( \mathbf{a} \). Indeed, if \( t \geq 0 \), then \( t \mathbf{a} \) has the same direction as \( \mathbf{a} \), but its magnitude, \( |t| \|\mathbf{a}\| \), is \( \geq \) or \( < \) the magnitude of \( \mathbf{a} \), depending on whether \( t \geq 1 \) or \( t < 1 \). When \( t \) is negative, \( t \mathbf{a} \) points in the opposite direction from \( \mathbf{a} \) but is again stretched or compressed depending on the size of \( |t| \).
Section 8.1  Algebraic Structure  271

Using $\mathbb{R}^2$ as a guide, we can extend concepts from $\mathbb{R}^2$ to $\mathbb{R}^n$. Here are five examples.

1) Every $(a, b) \in \mathbb{R}^2$ can be written as $(a, b) = a(1, 0) + b(0, 1)$. Using this as a guide, we define the usual basis of $\mathbb{R}^n$ to be the collection $\{e_1, \ldots, e_n\}$, where $e_j$ is the point in $\mathbb{R}^n$ whose $j$th coordinate is 1, and all other coordinates are 0. Notice by definition that each $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ can be written as a linear combination of the $e_j$’s:

$$x = \sum_{j=1}^{n} x_j e_j.$$

We shall not discuss other bases of $\mathbb{R}^n$ or the more general concept of “vector spaces,” which can be introduced using postulates similar in spirit to Postulate 1 in Chapter 1. Instead, we have introduced just enough algebraic machinery in $\mathbb{R}^n$ to develop the calculus of multivariable functions. For more information about $\mathbb{R}^n$ and abstract vector spaces, see Noble and Daniel [9].

Note: In $\mathbb{R}^2$ or $\mathbb{R}^3$, $e_1$ is denoted by $i$, $e_2$ is denoted by $j$, and, in $\mathbb{R}^3$, $e_3$ is denoted by $k$. Thus, in $\mathbb{R}^3$, $i := (1, 0, 0)$, $j := (0, 1, 0)$, and $k := (0, 0, 1)$.

2) Let $\phi : \mathbb{R} \rightarrow \mathbb{R}^2$ and $a, b \in \mathbb{R}^2$ with $b$ nonzero. By the geometric interpretation of vector addition, $\phi(t) := a + tb$ is a point on the line passing through $a$ in the direction of $b$. Using this as a guide, we define the straight line in $\mathbb{R}^n$ which passes through a point $a \in \mathbb{R}^n$ in the direction $b \in \mathbb{R}^n \setminus \{0\}$ to be the set of points

$$\ell_a(b) := \{a + tb : t \in \mathbb{R}\}.$$

In particular, it is easy to see that the parallelogram $\mathcal{P}(a, b)$ determined by nonzero vectors $a$ and $b$ in $\mathbb{R}^n$ can be described as

$$\mathcal{P}(a; b) := \{ua + vb : u, v \in [0, 1]\}.$$

3) Fix $a \neq b$ in $\mathbb{R}^2$, and set $\psi(t) := (1 - t)a + tb$, for $t \in \mathbb{R}$. Since $\psi(t) = a + t(b - a)$, it is evident that $\psi$ describes the line $\ell_a(b - a)$. This line passes through the
points $\psi(t) := a + tb$. In fact, by the geometric interpretation of vector subtraction, as $t$ ranges from 0 to 1, the points $\psi(t)$ trace out the diagonal of $P(a, b)$ that does not contain the origin (see Figure 8.1). It begins at $a$ and ends at $b$. Using this as a guide, we define the line segment from $a \in \mathbb{R}^n$ to $b \in \mathbb{R}^n$ to be the set of points

$$L(a; b) := \{(1 - t)a + tb : t \in [0, 1]\}.$$  

4) The angle between two nonzero vectors $a, b \in \mathbb{R}^2$ can be computed by the following process. If $\Delta$ is the triangle determined by the points $0, a,$ and $b,$ then the sides of $\Delta$ have length $\|a\|, \|b\|,$ and $\|a - b\|.$ If we let $\theta$ be the angle between $a$ and $b$ (i.e., the angle in $\Delta$ at the vertex $(0, 0)$), then by the Law of Cosines (see Appendix B),

$$\|a - b\|^2 = \|a\|^2 + \|b\|^2 - 2\|a\|\|b\| \cos \theta.$$  

Since Theorem 8.2 implies $\|a - b\|^2 = (a - b) \cdot (a - b) = \|a\|^2 - 2a \cdot b + \|b\|^2,$ it follows that $-2a \cdot b = -2\|a\|\|b\| \cos \theta.$ Since neither $a$ nor $b$ is zero, we conclude that

$$\cos \theta = \frac{a \cdot b}{\|a\|\|b\|}.$$  

(2)

Using this as a guide, we define the angle between two nonzero vectors $a, b \in \mathbb{R}^n$ (for any $n \in \mathbb{N}$) to be the number $\theta \in [0, \pi]$ determined by (2). (Our next result, the Cauchy–Schwarz Inequality, shows that the right side of (2) always belongs to the interval $[-1, 1].$ Hence, for each pair of nonzero vectors $a, b \in \mathbb{R}^n,$ there is a unique angle $\theta \in [0, \pi]$ which satisfies (2).)

5) Two vectors in $\mathbb{R}^2$ are parallel when one is a multiple of the other, and orthogonal when the angle, $\theta,$ between them is $\pi/2;$ that is, when $a \cdot b = \cos \theta \|a\|\|b\| = 0.$ Using this as a guide, we make the following definition in $\mathbb{R}^n$.

8.4 Definition.

Let $a$ and $b$ be nonzero vectors in $\mathbb{R}^n$.

i) $a$ and $b$ are said to be parallel if and only if there is a scalar $t \in \mathbb{R}$ such that $a = tb$.

ii) $a$ and $b$ are said to be orthogonal if and only if $a \cdot b = 0$.

Notice that the usual basis $(e_j)$ consists of pairwise orthogonal vectors; that is, $e_j \cdot e_k = 0$ when $j \neq k.$ In particular, the usual basis is an orthogonal basis.

We note in passing that Definition 8.4 is consistent with formula (2)—see Exercise 8.1.4b. Indeed, if $\theta$ is the angle between two nonzero vectors $a$ and $b$ in $\mathbb{R}^n,$ then $a$ and $b$ are parallel if and only if $\theta = 0$ or $\theta = \pi,$ and $a$ and $b$ are orthogonal if and only if $\theta = \pi/2.$

We shall see below that in addition to suggesting definitions for $\mathbb{R}^n,$ the geometry of $\mathbb{R}^2$ can also be used to help suggest proof strategies in $\mathbb{R}^n.$
Let's return to the analogy between $\mathbb{R}$ and $\mathbb{R}^n$. Surely, if we are going to develop a calculus of several variables, we need to know more about the Euclidean norm on $\mathbb{R}^n$. The next two results answer the question, How many properties do the absolute value and the Euclidean norm share?

Although the norm is not multiplicative, the following fundamental inequality can be used as a replacement for the multiplicative property in most proofs. (Some authors call this the Cauchy–Schwarz–Bunyakovsky Inequality.)

**8.5 Theorem.** [Cauchy–Schwarz Inequality].

If $x, y \in \mathbb{R}^n$, then

$$|x \cdot y| \leq \|x\| \|y\|.$$  

**Strategy:** Using the fact that the dot product of a vector with itself is the square of the norm of the vector and the square of any real number is nonnegative, identity (1) becomes $0 \leq \|x - y\|^2 = \|x\|^2 - 2x \cdot y + \|y\|^2$. We could solve this inequality to get an estimate of the dot product of $x \cdot y$, but this estimate might be very crude if $\|x - y\|$ were much larger than zero. But $x - y$ is only one point on the line $l_x(y)$. We might get a better estimate of the dot product $x \cdot y$ by using the inequality

$$0 \leq \|x - ty\|^2 = (x - ty) \cdot (x - ty) = \|x\|^2 - 2t(x \cdot y) + t^2\|y\|^2 \quad (3)$$

for other values of $t$. In fact, if we draw a picture in $\mathbb{R}^2$ (see Figure 8.2), we see that the norm of $\|x - ty\|$ is smallest for the value of $t$ which makes $x - ty$ orthogonal to $y$, that is, when

$$0 = (x - ty) \cdot y = x \cdot y - ty \cdot y = x \cdot y - t\|y\|^2.$$  

This suggests using $t = x \cdot y/\|y\|^2$ when $y \neq 0$. It turns out that this value of $t$ is exactly the one which reproduces the Cauchy–Schwarz Inequality. Here are the details.
Proof. The Cauchy–Schwarz Inequality is trivial when \( y = 0 \). If \( y \neq 0 \), substitute \( t = (x \cdot y)/\|y\|^2 \) into (3) to obtain

\[
0 \leq \|x\|^2 - t(x \cdot y) = \|x\|^2 - \frac{(x \cdot y)^2}{\|y\|^2}.
\]

It follows that \( 0 \leq \|x\|^2 - (x \cdot y)^2/\|y\|^2 \). Solving this inequality for \( (x \cdot y)^2 \), we conclude that

\[
(x \cdot y)^2 \leq \|x\|^2 \|y\|^2.
\]

The analogy between the absolute value and the Euclidean norm is further reinforced by the following result (compare with Theorem 1.7). (See also Exercise 8.1.10.)

8.6 Theorem. Let \( x, y \in \mathbb{R}^n \). Then

i) \( \|x\| \geq 0 \) with equality only when \( x = 0 \),

ii) \( \|ax\| = |a|\|x\| \) for all scalars \( a \),

iii) [Triangle Inequalities]. \( \|x + y\| \leq \|x\| + \|y\| \) and \( \|x - y\| \geq \|x\| - \|y\| \).

Proof. Statements i) and ii) are easy to verify.

To prove iii), observe that by Definition 8.3, Theorem 8.2, and the Cauchy–Schwarz Inequality,

\[
\|x + y\|^2 = (x + y) \cdot (x + y) = x \cdot x + 2x \cdot y + y \cdot y
= \|x\|^2 + 2x \cdot y + \|y\|^2 \leq \|x\|^2 + 2\|x\| \|y\| + \|y\|^2 = (\|x\| + \|y\|)^2.
\]

This establishes the first inequality in iii). By modifying the proof of Theorem 1.7, we can also establish the second inequality in iii).

Notice that the Triangle Inequality has a simple geometric interpretation. Indeed, since \( \|x\| \) is the magnitude of the vector \( x \), the inequality \( \|x + y\| \leq \|x\| + \|y\| \) states that the length of one side of a triangle (namely, the triangle whose vertices are \( 0, x, \) and \( x + y \)) is less than or equal to the sum of the lengths of its other two sides.

For some estimates, it is convenient to relate the Euclidean norm to the \( \ell^1 \)-norm and the sup-norm.

8.7 Remark. Let \( x \in \mathbb{R}^n \). Then

i) \( \|x\|_\infty \leq \|x\| \leq \sqrt{n} \|x\|_\infty \), and

ii) \( \|x\|_1 \leq \|x\| \leq \sqrt{n} \|x\|_1 \).
Section 8.1  Algebraic Structure  275

Proof. i) Let \( 1 \leq j \leq n \). By definition,

\[
|x_j|^2 \leq \|x\|^2 = x_1^2 + \cdots + x_n^2 \leq n \left( \max_{1 \leq \ell \leq n} |x_\ell| \right)^2 = n \|x\|^2_\infty;
\]

that is, \(|x_j| \leq \|x\|\) and \(\|x\| \leq \sqrt{n}\|x\|_\infty\). Taking the supremum of the first of these inequalities, over all \(1 \leq j \leq n\), we also have \(\|x\|_\infty \leq \|x\|\).

ii) Let \( A = \{(i, j) : 1 \leq i, j \leq n \text{ and } i < j\} \). To verify the first inequality, observe by algebra that

\[
\|x\|^2 = \left( \sum_{i=1}^{n} |x_i| \right)^2 = \sum_{i=1}^{n} |x_i|^2 + 2 \sum_{(i, j) \in A} |x_i| |x_j| = \|x\|^2 + 2 \sum_{(i, j) \in A} |x_i| |x_j|.
\]

Since \(\sum_{(i, j) \in A} |x_i| |x_j| \geq 0\), it follows that \(\|x\|^2 \leq \|x\|^2_1\).

On the other hand,

\[
0 \leq \sum_{(i, j) \in A} (|x_i| - |x_j|)^2 = \sum_{i=1}^{n} (n-1)|x_i|^2 - 2 \sum_{(i, j) \in A} |x_i| |x_j|
\]

\[
= n\|x\|^2 - \left( \sum_{i=1}^{n} |x_i|^2 + 2 \sum_{(i, j) \in A} |x_i| |x_j| \right) = n\|x\|^2 - \|x\|^2_1.
\]

This proves the second inequality.

Since \(x \cdot y\) is a scalar, the dot product in \(\mathbb{R}^n\) does not satisfy the closure property for any \(n > 1\). Here is another product, defined only on \(\mathbb{R}^2\), which does satisfy the closure property. (As we shall see below, this product allows us to exploit the geometry of \(\mathbb{R}^3\) in several unique ways.)

8.8 Definition.

The cross product of two vectors \(x = (x_1, x_2, x_3)\) and \(y = (y_1, y_2, y_3)\) in \(\mathbb{R}^3\) is the vector defined by

\[
x \times y := (x_2y_3 - x_3y_2, x_3y_1 - x_1y_3, x_1y_2 - x_2y_1).
\]

Using the usual basis \(i = e_1, j = e_2, k = e_3\), and the determinant operator (see Appendix C), we can give the cross product a more easily remembered form:

\[
x \times y = \det \begin{bmatrix} i & j & k \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{bmatrix}.
\]
The following result shows that the cross product satisfies some, but not all, of the usual laws of algebra. (Specifically, notice that although the cross product satisfies the distributive property, it satisfies neither the commutative property nor the associative property.)

8.9 Theorem. Let \( \mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^3 \) be vectors and \( \alpha \) be a scalar. Then

i) \[ \mathbf{x} \times \mathbf{x} = \mathbf{0}, \quad \mathbf{x} \times \mathbf{y} = -\mathbf{y} \times \mathbf{x}, \]

ii) \[ (\alpha \mathbf{x}) \times \mathbf{y} = \alpha (\mathbf{x} \times \mathbf{y}) = \mathbf{x} \times (\alpha \mathbf{y}), \]

iii) \[ \mathbf{x} \times (\mathbf{y} + \mathbf{z}) = (\mathbf{x} \times \mathbf{y}) + (\mathbf{x} \times \mathbf{z}), \]

iv) \[ (\mathbf{x} \times \mathbf{y}) \cdot \mathbf{z} = \mathbf{x} \cdot (\mathbf{y} \times \mathbf{z}) = \det \begin{bmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{bmatrix}, \]

v) \[ \mathbf{x} \times (\mathbf{y} \times \mathbf{z}) = (\mathbf{x} \cdot \mathbf{z})\mathbf{y} - (\mathbf{x} \cdot \mathbf{y})\mathbf{z}, \]

and

vi) \[ ||\mathbf{x} \times \mathbf{y}||^2 = (\mathbf{x} \cdot \mathbf{y})(\mathbf{y} \cdot \mathbf{y}) - (\mathbf{x} \cdot \mathbf{y})^2. \]

vii) Moreover, if \( \mathbf{x} \times \mathbf{y} \neq \mathbf{0} \), then the vector \( \mathbf{x} \times \mathbf{y} \) is orthogonal to \( \mathbf{x} \) and \( \mathbf{y} \).

Proof. These properties follow immediately from the definitions. We will prove properties iv), v), and vii) and leave the rest as an exercise.

iv) Notice that by definition,

\[
(\mathbf{x} \times \mathbf{y}) \cdot \mathbf{z} = (x_2y_3 - x_3y_2)x_1 + (x_3y_1 - x_1y_3)x_2 + (x_1y_2 - x_2y_1)x_3
= x_1(y_2z_3 - y_3z_2) + x_2(y_3z_1 - y_1z_3) + x_3(y_1z_2 - y_2z_1).
\]

Since this last expression is both the scalar \( \mathbf{x} \cdot (\mathbf{y} \times \mathbf{z}) \) and the value of the determinant on the right side of iv) (expanded along the first row), this verifies iv).

v) Since \( \mathbf{x} \times (\mathbf{y} \times \mathbf{z}) = (x_1, x_2, x_3) \times (y_2z_3 - y_3z_2, y_3z_1 - y_1z_3, y_1z_2 - y_2z_1) \), the first component of \( \mathbf{x} \times (\mathbf{y} \times \mathbf{z}) \) is

\[
x_2y_1z_2 - x_2y_2z_1 - x_3y_2z_1 + x_3y_3z_1 = (x_1z_1 + x_2z_2 + x_3z_3)y_1 - (x_1y_1 + x_2y_2 + x_3y_3)z_1.
\]

This proves that the first components of \( \mathbf{x} \times (\mathbf{y} \times \mathbf{z}) \) and \( (\mathbf{x} \cdot \mathbf{z})\mathbf{y} - (\mathbf{x} \cdot \mathbf{y})\mathbf{z} \) are equal. A similar argument shows that the second and third components are also equal.

vii) By parts i) and iv), \( (\mathbf{x} \times \mathbf{y}) \cdot \mathbf{x} = - (\mathbf{y} \times \mathbf{x}) \cdot \mathbf{x} = - \mathbf{y} \cdot (\mathbf{x} \times \mathbf{x}) = - \mathbf{y} \cdot \mathbf{0} = 0 \).

Thus \( \mathbf{x} \times \mathbf{y} \) is orthogonal to \( \mathbf{x} \). A similar calculation shows that \( \mathbf{x} \times \mathbf{y} \) is orthogonal to \( \mathbf{y} \).

Part vii) is illustrated in Figure 8.3. Notice that \( \mathbf{x} \times \mathbf{y} \) satisfies the "right-hand" rule. Indeed, if one puts the fingers of the right hand along \( \mathbf{x} \) and the palm of the right hand along \( \mathbf{y} \), then the thumb points in the direction of \( \mathbf{x} \times \mathbf{y} \).
By (2), there is a close connection between dot products and cosines. The following result shows that there is a similar connection between cross products and sines.

8.10 Remark. Let \( \mathbf{x}, \mathbf{y} \) be nonzero vectors in \( \mathbb{R}^3 \) and \( \theta \) be the angle between \( \mathbf{x} \) and \( \mathbf{y} \). Then

\[
\| \mathbf{x} \times \mathbf{y} \| = \| \mathbf{x} \| \| \mathbf{y} \| \sin \theta.
\]

Proof. By Theorem 8.9vi and (2),

\[
\| \mathbf{x} \times \mathbf{y} \|^2 = (\| \mathbf{x} \| \| \mathbf{y} \| \cos \theta)^2 - (\| \mathbf{x} \| \| \mathbf{y} \| \cos \theta)^2 \\
= (\| \mathbf{x} \| \| \mathbf{y} \|)^2(1 - \cos^2 \theta) = (\| \mathbf{x} \| \| \mathbf{y} \|)^2 \sin^2 \theta.
\]

This observation can be used to establish a connection between cross products and area or volume (see Exercise 8.2.7).

EXERCISES

8.1.1. Let \( x, y, z \in \mathbb{R}^n \).

a) If \( \| \mathbf{x} - \mathbf{z} \| < 2 \) and \( \| \mathbf{y} - \mathbf{z} \| < 3 \), prove that \( \| \mathbf{x} - \mathbf{y} \| < 5 \).

b) If \( \| \mathbf{x} \| < 2 \), \( \| \mathbf{y} \| < 3 \), and \( \| \mathbf{z} \| < 4 \), prove that \( |\mathbf{x} \cdot \mathbf{y} - \mathbf{x} \cdot \mathbf{z}| < 14 \).

c) If \( \| \mathbf{x} - \mathbf{y} \| < 2 \) and \( \| \mathbf{z} \| < 3 \), prove that \( |\mathbf{x} \cdot (\mathbf{y} - \mathbf{z}) - \mathbf{y} \cdot (\mathbf{x} - \mathbf{z})| < 6 \).

d) If \( \| \mathbf{x} - \mathbf{y} \| < 2 \) and \( \| \mathbf{y} \| < 1 \), prove that \( |\| \mathbf{x} - \mathbf{y} \| \|^2 - \mathbf{x} \cdot \mathbf{x}| < 2 \).

e) If \( n = 3 \), \( \| \mathbf{x} - \mathbf{y} \| < 2 \), and \( \| \mathbf{z} \| < 3 \), prove that \( \| \mathbf{x} \times \mathbf{y} - \mathbf{x} \times \mathbf{z} \| < 6 \).

f) If \( n = 3 \), \( \| \mathbf{x} \| < 1 \), \( \| \mathbf{y} \| < 2 \), and \( \| \mathbf{z} \| < 3 \), prove that \( \| \mathbf{x} \cdot (\mathbf{y} \times \mathbf{z}) \| < 6 \).
8.1.2. Let $B := \{ x \in \mathbb{R}^n : \|x\| \leq 1 \}$.

a) If $a, b, c \in B$ and

$$v := \frac{(a \cdot b)c + (a \cdot c)b + (c \cdot b)a}{3},$$

prove that $v$ belongs to $B$.

b) If $a, b \in B$, prove that

$$|a \cdot c - b \cdot d| \leq \|b - c\| + \|a - d\|$$

for all $c, d \in \mathbb{R}^n$.

c) If $a, b, c \in B$ and $n = 3$, prove that

$$\sqrt{|a \cdot (b \times c)|^2 + |a \cdot b|^2} \leq 1.$$

8.1.3. Use the proof of Theorem 8.5 to show that equality in the Cauchy-Schwarz Inequality holds if and only if $x = 0$, $y = 0$, or $x$ is parallel to $y$.

8.1.4. Let $a$ and $b$ be nonzero vectors in $\mathbb{R}^n$.

a) If $\phi(t) = a + tb$ for $t \in \mathbb{R}$, show that for each $t_0, t_1, t_2 \in \mathbb{R}$ with $t_1, t_2 \neq t_0$, the angle between $\phi(t_1) - \phi(t_0)$ and $\phi(t_2) - \phi(t_0)$ is $0$ or $\pi$.

b) If $\theta$ is the angle between $a$ and $b$, show that $a$ and $b$ are parallel according to Definition 8.4 if and only if $\theta = 0$ or $\pi$, and that $a$ and $b$ are orthogonal according to Definition 8.4 if and only if $\theta = \pi/2$.

8.1.5. The midpoint of a side of a triangle in $\mathbb{R}^3$ is the point that bisects that side (i.e., that divides it into two equal pieces). Let $\Delta$ be a triangle in $\mathbb{R}^3$ with sides $A, B$, and $C$ and let $L$ denote the line segment between the midpoints of $A$ and $B$. Prove that $L$ is parallel to $C$ and that the length of $L$ is one-half the length of $C$.

8.1.6. a) Prove that $(1, 2, 3), (4, 5, 6), (0, 4, 2)$ are vertices of a right triangle in $\mathbb{R}^3$.

b) Find all nonzero vectors orthogonal to $(1, -1, 0)$ which lie in the plane $z = x$.

c) Find all nonzero vectors orthogonal to the vector $(3, 2, -5)$ whose components sum to 4.

8.1.7. Let $a < b$ be real numbers. The Cartesian product $[a, b] \times [a, b]$ is obviously a square in $\mathbb{R}^2$. Define a cube $Q$ in $\mathbb{R}^n$ to be the $n$-fold Cartesian product of $[a, b]$ with itself; that is, $Q := [a, b] \times \cdots \times [a, b]$. Find a formula of the angle between the longest diagonal of $Q$ and any of its edges. Show that when $n = 3$, this angle is approximately 54.74 degrees.

8.1.8. a) Using Postulate 1 in Section 1.2 and Definition 8.1, prove Theorem 8.2.

b) Prove Theorem 8.9, parts i) through iii) and vi).

c) Prove that if $x, y \in \mathbb{R}^3$, then $\|x \times y\| \leq \|x\| \|y\|$. 

8.1.9. Suppose that \( \{a_k\} \) and \( \{b_k\} \) are sequences of real numbers which satisfy

\[
\sum_{k=1}^{\infty} a_k^2 < \infty \quad \text{and} \quad \sum_{k=1}^{\infty} b_k^2 < \infty.
\]

Prove that the infinite series \( \sum_{k=1}^{\infty} a_k b_k \) converges absolutely.

8.1.10. Prove that the \( \ell^1 \)-norm and the sup-norm also satisfy Theorem 8.6.

8.2 PLANES AND LINEAR TRANSFORMATIONS

A plane \( \Pi \) in \( \mathbb{R}^3 \) is a set of points that is "flat" in some sense. What do we mean by flat? Any vector that lies in \( \Pi \) is orthogonal to a common direction, called the normal, which we will denote by \( \mathbf{b} \). Fix a point \( \mathbf{a} \in \Pi \). Since the vector \( \mathbf{x} - \mathbf{a} \) lies in \( \Pi \) for all \( \mathbf{x} \in \Pi \), and since two vectors are orthogonal when their dot product is zero, we see that \( (\mathbf{x} - \mathbf{a}) \cdot \mathbf{b} = 0 \) for all \( \mathbf{x} \in \Pi \) (see Figure 8.4).

Using this three-dimensional case as a guide, for any \( \mathbf{a}, \mathbf{b} \in \mathbb{R}^n \) with \( \mathbf{b} \neq 0 \), we call the set

\[
\Pi_{\mathbf{b}}(\mathbf{a}) := \{ \mathbf{x} \in \mathbb{R}^n : (\mathbf{x} - \mathbf{a}) \cdot \mathbf{b} = 0 \}
\]

the hyperplane in \( \mathbb{R}^n \) passing through a point \( \mathbf{a} \in \mathbb{R}^n \) with normal \( \mathbf{b} \). (We call it a plane when \( n = 3 \).) In particular, \( \Pi_{\mathbf{b}}(\mathbf{a}) \) is the set of all points \( \mathbf{x} \) such that \( \mathbf{x} - \mathbf{a} \) is orthogonal to \( \mathbf{b} \).

There is nothing unique about "the normal" of a hyperplane. Any nonzero vector \( \mathbf{c} \) parallel to \( \mathbf{b} \) will define the same hyperplane. Indeed, if \( \mathbf{b} \) and \( \mathbf{c} \) are parallel, then, by definition, \( \mathbf{b} = t \mathbf{c} \) for some nonzero \( t \in \mathbb{R} \); hence \( (\mathbf{x} - \mathbf{a}) \cdot \mathbf{b} = 0 \) if and only if \( (\mathbf{x} - \mathbf{a}) \cdot \mathbf{c} = 0 \). Nevertheless, many properties of hyperplane can be

![Diagram of a hyperplane](image-url)
determined by their normals. For example, the *angle between two hyperplanes* with respective normals \( \mathbf{b} \) and \( \mathbf{c} \) is defined to be the angle between the normals \( \mathbf{b} \) and \( \mathbf{c} \).

By an *equation* of a hyperplane \( \mathcal{H} \), we mean an expression of the form \( F(x) = 0 \), where \( F : \mathbb{R}^n \to \mathbb{R} \) is a function determined by the following property: A point \( \mathbf{x} \) belongs to \( \mathcal{H} \) if and only if \( F(\mathbf{x}) = 0 \). By definition, then, an equation of the hyperplane \( \mathcal{H}_B(a) \) [i.e., the hyperplane passing through the point \( \mathbf{a} = (a_1, \ldots, a_n) \) with normal \( \mathbf{b} = (b_1, \ldots, b_n) \)] is given by

\[
\mathbf{b} \cdot \mathbf{x} = \mathbf{b} \cdot \mathbf{a}.
\]

This form is sometimes referred to as the *point-normal form*. It can also be written in the form

\[
b_1x_1 + b_2x_2 + \cdots + b_nx_n = d,
\]

where \( d = b_1a_1 + b_2a_2 + \cdots + b_na_n \) is a constant determined by \( \mathbf{a} \) and \( \mathbf{b} \) (and related to the distance from \( \mathcal{H}_B(a) \) to the origin—see Exercise 8.2.8). In particular, planes in \( \mathbb{R}^3 \) have equations of the form

\[
ax + by + cz = d.
\]

Notice that a "hyperplane" in \( \mathbb{R}^2 \) is by definition a straight line. Just as straight lines through the origin played a prominent role in characterizing differentiability of functions of one variable (see Theorem 4.3), even so hyperplane-like objects will play a crucial role in defining differentiability of functions of several variables. Why hyperplane-like objects and not just hyperplanes themselves? Equations of hyperplanes are by definition real valued and we do not want to restrict our analysis of differentiable functions to the real-valued case.

What kind of hyperplane-like objects will be rich enough to develop a general theory for differentiability of vector-valued functions? To answer this question, we make the following observation about equations of straight lines through the origin. (Here we use \( s \) for slope since \( m \) will be used for the dimension of the range space \( \mathbb{R}^m \).)

**8.11 Remark.** Let \( T : \mathbb{R} \to \mathbb{R} \). Then \( T(x) = sx \) for some \( s \in \mathbb{R} \) if and only if \( T \) satisfies

\[
T(x + y) = T(x) + T(y) \quad \text{and} \quad T(\alpha x) = \alpha T(x)
\]

for all \( x, y, \alpha \in \mathbb{R} \).

**Proof.** If \( T(x) = sx \), then \( T \) satisfies (4) since the distributive and commutative laws hold on \( \mathbb{R} \). Conversely, if \( T \) satisfies (4), set \( s := T(1) \). Then (let \( \alpha = x \)),

\[
T(x) = T(x \cdot 1) = xT(1) = sx
\]

for all \( x \in \mathbb{R} \).
Accordingly, we introduce the following concept.

**8.12 Definition.**

A function \( T : \mathbb{R}^n \to \mathbb{R}^m \) is said to be **linear** [notation: \( T \in \mathcal{L}(\mathbb{R}^n; \mathbb{R}^m) \)] if and only if it satisfies

\[
T(x + y) = T(x) + T(y) \quad \text{and} \quad T(\alpha x) = \alpha T(x)
\]

for all \( x, y \in \mathbb{R}^n \) and all scalars \( \alpha \).

When \( m = 1 \) (i.e., when the range of \( T \) is \( \mathbb{R} \)), we shall often drop the boldface notation (i.e., write \( T \) for \( T \)).

Notice once and for all that if \( T \) is a linear function, then

\[
T(0) = 0.
\]  \hspace{1cm} (5)

Indeed, by definition, \( T(0) = T(0 + 0) = T(0) + T(0) \). Hence (5) can be obtained by subtracting \( T(0) \) from both sides of this last equation. Also notice that if \( F(x) = 0 \) is the equation of a hyperplane passing through the origin, then \( F(x) = a_1 x_1 + \cdots + a_n x_n \). In particular, \( F \in \mathcal{L}(\mathbb{R}^n; \mathbb{R}) \).

Functions in \( \mathcal{L}(\mathbb{R}^n; \mathbb{R}^m) \) are sometimes called **linear transformations** or **linear operators** because of the fundamental role they play in the theory of change of variables in \( \mathbb{R}^n \). We shall take up this connection in Chapter 12.

According to Remark 8.11, linear transformations of one variable [i.e., objects \( T \in \mathcal{L}(\mathbb{R}; \mathbb{R}) \)] can be identified with \( \mathbb{R} \) by representing \( T \) by its slope \( s \). Is there an analogue of slope which can be used to represent linear transformations of several variables? To answer this question, we use the following half page to review some elementary linear algebra.

Recall that an \( m \times n \) **matrix** \( B \) is a rectangular array which has \( m \) rows and \( n \) columns:

\[
B = [b_{ij}]_{m \times n} := \begin{bmatrix}
  b_{11} & b_{12} & \cdots & b_{1n} \\
  b_{21} & b_{22} & \cdots & b_{2n} \\
  \vdots & \vdots & \ddots & \vdots \\
  b_{m1} & b_{m2} & \cdots & b_{mn}
\end{bmatrix}
\]

For us, the **entries** \( b_{ij} \) of a matrix \( B \) will usually be numbers or real-valued functions. Let \( B = [b_{ij}]_{m \times n} \) and \( C = [c_{jk}]_{p \times q} \) be such matrices. Recall that the **product** of \( B \) and a scalar \( \alpha \) is defined by

\[
\alpha B = [\alpha b_{ij}]_{m \times n}.
\]
the sum of \( B \) and \( C \) is defined (when \( m = p \) and \( n = q \)) by

\[
B + C = [b_{ij} + c_{ij}]_{m \times n},
\]

and the product of \( B \) and \( C \) is defined (when \( n = p \)) by

\[
BC = \left[ \sum_{\nu=1}^{n} b_{\nu j} c_{\nu i} \right]_{m \times q}.
\]

Also recall that most of the usual laws of algebra hold for addition and multiplication of matrices (see Theorem C.1 in Appendix C). One glaring exception is that matrix multiplication is not commutative.

We shall identify points \( x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n \) with \( 1 \times n \) row matrices or \( n \times 1 \) column matrices by setting

\[
[x] = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad \text{or} \quad [x] = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}^T := \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix},
\]

where \( B^T \) represents the transpose of a matrix \( B \) (see Appendix C). Abusing the notation slightly, we shall usually represent the product of an \( m \times n \) matrix \( B \) and an \( n \times 1 \) column matrix \([x]\) by \( Bx \). This notation is justified, as the following result shows, since the function \( x \rightarrow [x] \) takes vector addition to matrix addition, the dot product to matrix multiplication, and scalar multiplication to scalar multiplication.

**8.13 Remark.** If \( x, y \in \mathbb{R}^n \) and \( \alpha \) is a scalar, then

\[
[x + y] = [x] + [y], \quad [x \cdot y] = [x][y]^T, \quad \text{and} \quad [\alpha x] = \alpha [x].
\]

**Proof.** These laws follow immediately from the definitions of addition and multiplication of matrices and vectors. For example,

\[
[x + y] = [x_1 + y_1 \ x_2 + y_2 \ \cdots \ x_n + y_n] \\
= [x_1 \ x_2 \ \cdots \ x_n] + [y_1 \ y_2 \ \cdots \ y_n] = [x] + [y].
\]

The following result shows that each \( m \times n \) matrix gives rise to a linear function from \( \mathbb{R}^n \) to \( \mathbb{R}^m \).

**8.14 Remark.** Let \( B = [b_{ij}] \) be an \( m \times n \) matrix whose entries are real numbers and let \( e_1, \ldots, e_n \) represent the usual basis of \( \mathbb{R}^n \). If

\[
T(x) = Bx, \quad x \in \mathbb{R}^n,
\]

(6)
then $T$ is a linear function from $\mathbb{R}^n$ to $\mathbb{R}^m$ and the $j$th column of $B$ can be obtained by evaluating $T$ at $e_j$:

$$(b_{1j}, b_{2j}, \ldots, b_{mj}) = T(e_j), \quad j = 1, 2, \ldots, n. \quad (7)$$

**Proof.** Notice, first, that (7) holds by (6) and the definition of matrix multiplication. Next, observe by Remark 8.13 and the distributive law of matrix multiplication (see Theorem C.1) that

$$T(x + y) = B(x + y) = B([x] + [y]) = B[x] + B[y] = T(x) + T(y)$$

for all $x, y \in \mathbb{R}^n$. Similarly, $T(\alpha x) = B(\alpha x) = B(\alpha [x]) = \alpha B[x] = \alpha T(x)$ for all $x \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$. Thus $T \in \mathcal{L}(\mathbb{R}^n; \mathbb{R}^m)$. □

Remark 8.14 would barely be worth mentioning were it not the case that ALL linear functions from $\mathbb{R}^n$ to $\mathbb{R}^m$ have this form. Here, then, is the multidimensional analogue of Remark 8.11.

8.15 Theorem. For each $T \in \mathcal{L}(\mathbb{R}^n; \mathbb{R}^m)$ there is a matrix $B = [b_{ij}]_{m \times n}$ such that (6) holds. Moreover, the matrix $B$ is unique. Specifically, for each fixed $T$ there is only one $B$ which satisfies (6), and the columns of $B$ are defined by (7).

**Proof.** Uniqueness has been established in Remark 8.14. To prove existence, suppose that $T \in \mathcal{L}(\mathbb{R}^n; \mathbb{R}^m)$. Define $B$ by (7). Then

$$T(x) = T \left( \sum_{j=1}^{n} x_{j}e_{j} \right)$$

$$= \sum_{j=1}^{n} x_{j}T(e_{j}) = \sum_{j=1}^{n} x_{j}(b_{1j}, b_{2j}, \ldots, b_{mj})$$

$$= \left( \sum_{j=1}^{n} x_{j}b_{1j}, \sum_{j=1}^{n} x_{j}b_{2j}, \ldots, \sum_{j=1}^{n} x_{j}b_{mj} \right) = Bx. \quad \square$$

The unique matrix $B$ which satisfies (6) is called the matrix which represents $T$.

In Chapter 11 we shall use this point of view to define what it means for a function from $\mathbb{R}^n$ into $\mathbb{R}^m$ to be differentiable. At that point, we shall show that many of the one-dimensional results about differentiation remain valid in the multidimensional setting. Since the one-dimensional theory relied on estimates using the absolute values of various functions, we expect the theory in $\mathbb{R}^n$ to rely on estimates using the norms of various functions. Since some of those functions will be linear, the following concept will be useful in this regard.
8.16 Definition. Let \( T \in \mathcal{L}(\mathbb{R}^n; \mathbb{R}^m) \). The operator norm of \( T \) is the extended real number

\[
\|T\| := \sup_{\|x\| \neq 0} \frac{\|T(x)\|}{\|x\|}.
\]

One interesting corollary of Theorem 8.15 is that the operator norm of a linear function is always finite.

8.17 Theorem. Let \( T \in \mathcal{L}(\mathbb{R}^n; \mathbb{R}^m) \). Then the operator norm of \( T \) is finite and satisfies

\[
\|T(x)\| \leq \|T\| \|x\|
\]

for all \( x \in \mathbb{R}^n \).

Proof. Since \( T(0) = 0 \), (8) holds for \( x = 0 \). On the other hand, by Definition 8.16, (8) holds for \( x \neq 0 \). It remains to prove that the extended real number \( \|T\| \) is finite.

Let \( B \) be the \( m \times n \) matrix which represents \( T \), and suppose that the rows of \( T \) are given by \( b_1, \ldots, b_m \). By the definition of matrix multiplication and our identification of \( \mathbb{R}^m \) with \( m \times 1 \) matrices,

\[
T(x) = (b_1 \cdot x, \ldots, b_m \cdot x).
\]

If \( B = O \), then \( \|T\| = 0 \) and (8) is an equality. If \( B \neq O \), then, by the Cauchy–Schwarz Inequality, the square of the Euclidean norm of \( T(x) \) satisfies

\[
\|T(x)\|^2 = (b_1 \cdot x)^2 + \cdots + (b_m \cdot x)^2 \\
\leq (\|b_1\| \|x\|)^2 + \cdots + (\|b_m\| \|x\|)^2 \\
\leq m \cdot \max\{\|b_j\|^2 : 1 \leq j \leq m\} \|x\|^2 =: C \|x\|^2.
\]

Therefore, the quotients \( \|T(x)\|/\|x\| \) are bounded (by \( \sqrt{C} \)). It follows from the Completeness Axiom that \( \|T\| \) exists and is finite.

Theorem 8.17, an analogue of the Cauchy–Schwarz Inequality, will be used to estimate differentiable functions of several variables. If \( B \) is the matrix which represents a linear transformation \( T \), we will refer to the number \( \|T\| \) as the operator norm of \( B \), and denote it by \( \|B\| \). (For two other ways to calculate this norm, see Exercise 8.2.11.)

We close this section with an optional result which shows that under the identification of linear functions with matrices, function composition is taken to matrix multiplication. This, in fact, is why matrix multiplication is defined the way it is.
8.18 Remark. If \( T : \mathbb{R}^n \rightarrow \mathbb{R}^m \) and \( U : \mathbb{R}^m \rightarrow \mathbb{R}^p \) are linear, then so is \( U \circ T \). In fact, if \( B \) is the \( m \times n \) matrix which represents \( T \), and \( C \) is the \( p \times m \) matrix which represents \( U \), then \( CB \) is the matrix which represents \( U \circ T \).

**Proof.** Let \( e_1, \ldots, e_n \) be the usual basis of \( \mathbb{R}^n \), \( u_1, \ldots, u_m \) be the usual basis of \( \mathbb{R}^m \), and \( w_1, \ldots, w_p \) be the usual basis of \( \mathbb{R}^p \). If \( B = [b_{ij}]_{m \times n} \) represents \( T \) and \( C = [c_{jk}]_{p \times m} \) represents \( U \), then, by Theorem 8.15,

\[
\sum_{k=1}^{m} b_{kj} u_k = (b_{1j}, \ldots, b_{mj}) = T(e_j), \quad j = 1, 2, \ldots, n,
\]

and

\[
\sum_{v=1}^{p} c_{vk} w_v = (c_{1k}, \ldots, c_{pk}) = U(u_k), \quad k = 1, 2, \ldots, m.
\]

Hence

\[
(U \circ T)(e_j) = U(T(e_j)) = U \left( \sum_{k=1}^{m} b_{kj} u_k \right) = \sum_{k=1}^{m} b_{kj} U(u_k)
\]

\[
= \sum_{k=1}^{m} \sum_{v=1}^{p} b_{kj} c_{vk} w_v = \left( \sum_{k=1}^{m} b_{kj} c_{1k}, \ldots, \sum_{k=1}^{m} b_{kj} c_{pk} \right)
\]

for each \( 1 \leq j \leq n \). Since this last vector is the \( j \)-th column of the matrix \( CB \), it follows that \( CB \) is the matrix which represents \( U \circ T \).

**EXERCISES**

8.2.1. Let \( a, b, c \in \mathbb{R}^3 \).

a) Prove that if \( a, b, \) and \( c \) do not all lie on the same line, then an equation of the plane through these points is given by \((x, y, z) \cdot d = a \cdot d\), where

\[
d := (a - b) \times (a - c).
\]

b) Prove that if \( c \) does not lie on the line \( \phi(t) = t a + b, \ t \in \mathbb{R} \), then an equation of the plane that contains this line and the point \( c \) is given by \((x, y, z) \cdot d = b \cdot d\), where \( d := a \times (b - c) \).

8.2.2. a) Find an equation of the hyperplane through the points \( (1, 0, 0, 0), (2, 1, 0, 0), (0, 1, 1, 0), \) and \((0, 4, 0, 1) \).

b) Find an equation of the hyperplane that contains the lines \( \phi(t) = (t, t, 1, t), \ t \in \mathbb{R} \) and \( \psi(t) = (1, t, 1 + t, t), \ t \in \mathbb{R} \).

c) Find an equation of the plane parallel to the hyperplane \( x_1 + \cdots + x_n = \pi \) passing through the point \((1, 2, \ldots, n)\).
8.2.3. Find two lines in $\mathbb{R}^3$ which are not parallel but do not intersect.

8.2.4. Suppose that $T \in \mathcal{L}(\mathbb{R}^n; \mathbb{R}^m)$ for some $n, m \in \mathbb{N}$.
   a) Find the matrix representative of $T$ if $T(x, y, z, w) = (0, x + y, x - z, x + y + w)$.
   b) Find the matrix representative of $T$ if $T(x, y, z) = x - y + z$.
   c) Find the matrix representative of $T$ if $T(x_1, x_2, \ldots, x_n) = (x_1 - x_n, x_n - x_1)$.

8.2.5. Suppose that $T \in \mathcal{L}(\mathbb{R}^n; \mathbb{R}^m)$ for some $n, m \in \mathbb{N}$.
   a) If $T(1, 1) = (3, \pi, 0)$ and $T(0, 1) = (4, 0, 1)$, find the matrix representative of $T$.
   b) If $T(1, 1, 0) = (e, \pi, 1)$, $T(0, -1, 1) = (1, 0)$, and $T(1, 1, -1) = (1, 2)$, find the matrix representative of $T$.
   c) If $T(0, 1, 1, 0) = (3, 5, \pi)$, $T(0, 1, -1, 0) = (5, 3, \pi)$, and $T(0, 0, 0, -1) = (\pi, 3, 1)$, find all possible matrix representatives of $T$.
   d) If $T(1, 1, 0, 0) = (5, 4, 1, 0)$, $T(0, 1, 0, 0) = (1, 2, 0, 1)$, and $T(0, 0, 0, -1) = (\pi, 3, -1, 0)$, find all possible matrix representatives of $T$.

8.2.6. Suppose that $a, b, c \in \mathbb{R}^3$ are three points which do not lie on the same straight line and that $\Pi$ is the plane which contains the points $a, b, c$.
Prove that an equation of $\Pi$ is given by

$$
\det \begin{bmatrix}
x - a_1 & y - a_2 & z - a_3 \\
b_1 - a_1 & b_2 - a_2 & b_3 - a_3 \\
c_1 - a_1 & c_2 - a_2 & c_3 - a_3
\end{bmatrix} = 0.
$$

8.2.7. This exercise is used in Appendix E. Recall that the area of a parallelogram with base $b$ and altitude $h$ is given by $bh$, and the volume of a parallelepiped is given by the area of its base times its altitude.

   a) Let $a, b \in \mathbb{R}^3$ be nonzero vectors and $\mathcal{P}$ represent the parallelogram

$$
\{(x, y, z) = ua + vb : u, v \in [0, 1]\}.
$$

Prove that the area of $\mathcal{P}$ is $\|a \times b\|$.

   b) Let $a, b, c \in \mathbb{R}^3$ be nonzero vectors and $\mathcal{P}$ represent the parallelepiped

$$
\{(x, y, z) = ta + ub + vc : t, u, v \in [0, 1]\}.
$$

Prove that the volume of $\mathcal{P}$ is $\| (a \times b) \cdot c \|$.

8.2.8. The distance from a point $x_0 = (x_0, y_0, z_0)$ to a plane $\Pi$ in $\mathbb{R}^3$ is defined to be

$$
dist (x_0, \Pi) := \begin{cases}
0 & x_0 \in \Pi \\
\|v\| & x_0 \notin \Pi,
\end{cases}
$$
where \( \mathbf{v} := (x_0 - x_1, y_0 - y_1, z_0 - z_1) \) for some \((x_1, y_1, z_1) \in \mathcal{P}\), and \( \mathbf{v} \) is orthogonal to \( \mathcal{P} \) (i.e., parallel to its normal). Sketch \( \mathcal{P} \) and \( x_0 \) for a typical plane \( \mathcal{P} \), and convince yourself that this is the correct definition. Prove that this definition does not depend on the choice of \( \mathbf{v} \), by showing that the distance from \( x_0 = (x_0, y_0, z_0) \) to the plane \( \mathcal{P} \) described by \( ax + by + cz = d \) is

\[
\text{dist}(x_0, \mathcal{P}) = \frac{|ax_0 + by_0 + cz_0 - d|}{\sqrt{a^2 + b^2 + c^2}}.
\]

**8.2.9.** [Rotations in \( \mathbb{R}^2 \)]. This exercise is used in Section **15.1**. Let

\[
- B = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}
\]

for some \( \theta \in \mathbb{R} \).

a) Prove that \( \|B(x, y)\| = \|(x, y)\| \) for all \((x, y) \in \mathbb{R}^2\).

b) Let \((x, y) \in \mathbb{R}^2 \) be a nonzero vector and \( \varphi \) represent the angle between \( B(x, y) \) and \((x, y)\). Prove that \( \cos \varphi = \cos \theta \). Thus, show that \( B \) rotates \( \mathbb{R}^2 \) through an angle \( \theta \). (When \( \theta > 0 \), we shall call \( B \) counterclockwise rotation about the origin through the angle \( \theta \).)

**8.2.10.** For each of the following functions \( f \), find the matrix representative of a linear transformation \( T \in \mathcal{L}(\mathbb{R}, \mathbb{R}^m) \) which satisfies

\[
\lim_{h \to 0} \frac{\|T(x + h) - f(x) - T(h)\|}{h} = 0.
\]

a) \( f(x) = (x^2, \sin x) \)

b) \( f(x) = (e^x, \sqrt{x}, 1 - x^2) \)

c) \( f(x) = (1, 2, 3, x^2 + x, x^2 - x) \)

**8.2.11.** Fix \( T \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m) \). Set

\[
M_1 := \sup_{\|x\| = 1} \|T(x)\| \quad \text{and} \quad M_2 := \inf\{C > 0 : \|T(x)\| \leq C\|x\| \text{ for all } x \in \mathbb{R}^n\}.
\]

a) Prove that \( M_1 \leq \|T\| \).

b) Using the linear property of \( T \), prove that if \( x \neq 0 \), then

\[
\frac{\|T(x)\|}{\|x\|} \leq M_1.
\]

c) Prove that \( M_1 = M_2 = \|T\| \).
8.3 TOPOLOGY OF $\mathbb{R}^n$

If you want a more abstract introduction to the topology of Euclidean spaces, skip the rest of this chapter and the next, and begin Chapter 10 now.

Topology, a study of geometric objects which emphasizes how they are put together over their exact shape and proportion, is based on the fundamental concepts of open and closed sets, a generalization of open and closed intervals. In this section we introduce these concepts in $\mathbb{R}^n$ and identify their most basic properties. In the next chapter, we shall explore how they can be used to characterize limits and continuity without using distance explicitly. This additional step in abstraction will yield powerful benefits, as we shall see in Section 9.4 and in Chapter 11 when we begin to study the calculus of functions of several variables.

We begin with a natural generalization of intervals to $\mathbb{R}^n$.

8.19 Definition.

Let $a \in \mathbb{R}^n$.

i) For each $r > 0$, the open ball centered at $a$ of radius $r$ is the set of points

$$B_r(a) := \{ x \in \mathbb{R}^n : \|x - a\| < r \}. $$

ii) For each $r \geq 0$, the closed ball centered at $a$ of radius $r$ is the set of points

$$\{ x \in \mathbb{R}^n : \|x - a\| \leq r \}.$$

Notice that when $n = 1$, the open ball centered at $a$ of radius $r$ is the open interval $(a - r, a + r)$, and the corresponding closed ball is the closed interval $[a - r, a + r]$. Also notice that the open ball (respectively, the closed ball) centered at $a$ of radius $r$ contains none of its (respectively, all of its) "boundary" $\{ x : \|x - a\| = r \}$. Accordingly, we will draw pictures of balls in $\mathbb{R}^2$ with the following conventions: Open balls will be drawn with dashed "boundaries" and closed balls will be drawn with solid "boundaries" (see Figure 8.5).

To generalize the concept of open and closed intervals even further, observe that each element of an open interval $I$ lies "inside" $I$ (i.e., is surrounded by other points in $I$). On the other hand, although closed intervals do NOT satisfy this property, their complements do. Accordingly, we make the following definition.

8.20 Definition.

Let $n \in \mathbb{N}$.

i) A subset $V$ of $\mathbb{R}^n$ is said to be open (in $\mathbb{R}^n$) if and only if for every $a \in V$ there is an $\varepsilon > 0$ such that $B_\varepsilon(a) \subseteq V$.

ii) A subset $E$ of $\mathbb{R}^n$ is said to be closed (in $\mathbb{R}^n$) if and only if $E^c := \mathbb{R}^n \setminus E$ is open.
The following result shows that every "open" ball is open. (Closed balls are also closed—see Exercise 8.3.2.)

**8.21 Remark.** For every \( x \in B_r(a) \) there is an \( \varepsilon > 0 \) such that \( B_\varepsilon(x) \subseteq B_r(a) \).

**Proof.** Let \( x \in B_r(a) \). Using Figure 8.5 for guidance, we set \( \varepsilon = r - \|x - a\| \). If \( y \in B_\varepsilon(x) \), then by the Triangle Inequality, assumption, and the choice of \( \varepsilon \),
\[
\|y - a\| \leq \|y - x\| + \|x - a\| < \varepsilon + \|x - a\| = r.
\]
Thus, by definition, \( y \in B_r(a) \). In particular, \( B_\varepsilon(x) \subseteq B_r(a) \).

(Once again, drawing diagrams in \( \mathbb{R}^2 \) led us to a proof valid for all Euclidean spaces.)

Here are more examples of open sets and closed sets.

**8.22 Remark.** If \( a \in \mathbb{R}^n \), then \( \mathbb{R}^n \setminus \{a\} \) is open and \( \{a\} \) is closed.

**Proof.** By Definition 8.20, it suffices to prove that the complement of every singleton \( E := \{a\} \) is open. Let \( x \in E^c \) and set \( \varepsilon = \|x - a\| \). Then, by definition, \( a \notin B_\varepsilon(x) \), so \( B_\varepsilon(x) \subseteq E^c \). Therefore, \( E^c \) is open by Definition 8.20.

Students sometimes mistakenly believe that every set is either open or closed. Some sets are neither open nor closed (like the interval \([0, 1])\). And, as the
following result shows; every Euclidean space contains two special sets which are both open and closed. (We shall see below that these are the only subsets of \( \mathbb{R}^n \) which are simultaneously open and closed.)

**8.23 Remark.** For each \( n \in \mathbb{N} \), the empty set \( \emptyset \) and the whole space \( \mathbb{R}^n \) are both open and closed.

**Proof.** Since \( \mathbb{R}^n = \emptyset^c \) and \( \emptyset = (\mathbb{R}^n)^c \), it suffices by Definition 8.20 to prove that \( \emptyset \) and \( \mathbb{R}^n \) are both open. Because the empty set contains no points, "every" point \( x \in \emptyset \) satisfies \( B_\varepsilon(x) \subseteq \emptyset \). (This is called the vacuous implication.) Therefore, \( \emptyset \) is open. On the other hand, since \( B_\varepsilon(x) \subseteq \mathbb{R}^n \) for all \( x \in \mathbb{R}^n \) and all \( \varepsilon > 0 \), it is clear that \( \mathbb{R}^n \) is open.

It is important to recognize that open sets and closed sets behave very differently with respect to unions and intersections. (In fact, these properties are so important that they form the basis of an axiomatic system which describes all topological spaces, even those for which measurement of distance is impossible.)

**8.24 Theorem.** Let \( n \in \mathbb{N} \).

i) If \( \{V_a\}_{a \in A} \) is any collection of open subsets of \( \mathbb{R}^n \), then

\[
\bigcup_{a \in A} V_a
\]

is open.

ii) If \( \{V_k : k = 1, 2, \ldots, p\} \) is a finite collection of open subsets of \( \mathbb{R}^n \), then

\[
\bigcap_{k=1}^{p} V_k := \bigcap_{k \in \{1, 2, \ldots, p\}} V_k
\]

is open.

iii) If \( \{E_a\}_{a \in A} \) is any collection of closed subsets of \( \mathbb{R}^n \), then

\[
\bigcap_{a \in A} E_a
\]

is closed.

iv) If \( \{E_k : k = 1, 2, \ldots, p\} \) is a finite collection of closed subsets of \( \mathbb{R}^n \), then

\[
\bigcup_{k=1}^{p} E_k := \bigcup_{k \in \{1, 2, \ldots, p\}} E_k
\]

is closed.

v) If \( V \) is open and \( E \) is closed, then \( V \setminus E \) is open and \( E \setminus V \) is closed.
Section 8.3 Topology of \( \mathbb{R}^n \) 291

**Proof.** i) Let \( x \in \bigcup_{\alpha \in A} V_{\alpha} \). Then \( x \in V_{\alpha} \) for some \( \alpha \in A \). Since \( V_{\alpha} \) is open, it follows that there is an \( r > 0 \) such that \( B_r(x) \subseteq V_{\alpha} \). Thus \( B_r(x) \subseteq \bigcup_{\alpha \in A} V_{\alpha} \); that is, this union is open.

ii) Let \( x \in \bigcap_{k=1}^{p} V_k \). Then \( x \in V_k \) for \( k = 1, 2, \ldots, p \). Since each \( V_k \) is open, it follows that there are numbers \( r_k > 0 \) such that \( B_{r_k}(x) \subseteq V_k \). Let \( r = \min\{r_1, \ldots, r_p\} \). Then \( r > 0 \) and \( B_r(x) \subseteq V_k \) for all \( k = 1, 2, \ldots, p \); that is, \( B_r(x) \subseteq \bigcap_{k=1}^{p} V_k \). Hence, this intersection is open.

iii) By DeMorgan's Law (Theorem 1.36) and part i),

\[
\left( \bigcap_{\alpha \in A} E_{\alpha} \right)^c = \bigcup_{\alpha \in A} E_{\alpha}^c
\]

is open, so \( \bigcap_{\alpha \in A} E_{\alpha} \) is closed.

iv) By DeMorgan's Law and part ii),

\[
\left( \bigcup_{k=1}^{p} E_k \right)^c = \bigcap_{k=1}^{p} E_k^c
\]

is open, so \( \bigcup_{k=1}^{p} E_k \) is closed.

v) Since \( V \setminus E = V \cap E^c \) and \( E \setminus V = E \cap V^c \), the former is open by part ii), and the latter is closed by part iii).

The finiteness hypotheses in Theorem 8.24 are crucial, even for the case \( n = 1 \).

**8.25 Remark.** Statements ii) and iv) of Theorem 8.24 are false if arbitrary collections are used in place of finite collections.

**Proof.** In the Euclidean space \( \mathbb{R} \),

\[
\bigcap_{k \in \mathbb{N}} \left( -\frac{1}{k}, \frac{1}{k} \right) = \{0\}
\]

is closed and

\[
\bigcup_{k \in \mathbb{N}} \left[ \frac{1}{k+1}, \frac{k}{k+1} \right] = (0, 1)
\]

is open.

To see why open sets are so important to analysis, we reexamine the definition of continuity using open sets. By Definition 3.19, a function \( f : E \to \mathbb{R} \) is continuous at \( a \in E \) if and only if given \( \varepsilon > 0 \) there is a \( \delta > 0 \) such that \( |x-a| < \delta \) and \( x \in E \) imply \( |f(x) - f(a)| < \varepsilon \). Put in "ball language," this says that \( f \) is continuous at \( a \in E \) if and only if \( f(E \cap B_{\delta}(a)) \subseteq B_{\varepsilon}(f(a)) \); that is, \( E \cap B_{\delta}(a) \subseteq f^{-1}(B_{\varepsilon}(f(a))) \). In particular, \( f \) is continuous at \( a \in E \) if and only if for all
$a \in \mathbb{R}$, the inverse image under $f$ of every open ball centered at $f(a)$ contains the intersection of $E$ and an open ball centered at $a$.

Intersecting the ball centered at $a$ with $E$ adds a complication. It would be simpler if the inverse image of an open ball under a continuous function just contained another open ball. Can we discard the set $E$ like that? To answer this question, we consider two functions, $f(x) = 1/x$ and $g(x) = 1 + \sqrt{x - 1}$, and one open ball, $(-1, 3)$, centered at 1. Notice that $f^{−1}(-1, 3) = (−\infty, −1) \cup (1/3, \infty)$ does contain an open ball centered at $a = 1$ but $g^{−1}(-1, 3) = [1, 5)$ does not. It is merely the intersection of an open ball and the domain of $g$:

$g^{−1}(-1, 3) = [1, 5) = [1, \infty) \cap (−5, 5)$. Evidently, we cannot discard the domain $E$ of a continuous function when restating the definition of continuity using open balls, unless $E$ is open (see Exercise 9.4.3).

Accordingly, we modify the definition of open and closed along the following lines.

### 8.26 Definition.

Let $E \subseteq \mathbb{R}^n$.

i) A set $U \subseteq E$ is said to be relatively open in $E$ if and only if there is an open set $A$ such that $U = E \cap A$.

ii) A set $C \subseteq E$ is said to be relatively closed in $E$ if and only if there is a closed set $B$ such that $C = E \cap B$.

We postpone using these concepts to study continuous functions on $\mathbb{R}^n$ until Chapter 9. Meanwhile, we shall use relatively open sets to introduce connectivity, a concept which generalizes to $\mathbb{R}^n$ an important property of intervals which played a role in the proof of the Intermediate Value Theorem, and which will be used several times in our development of the calculus of functions of several variables. First, we explore the analogy between relatively open sets and open sets.

### 8.27 Remark. Let $U \subseteq E \subseteq \mathbb{R}^n$.

i) Then $U$ is relatively open in $E$ if and only if for each $a \in U$ there is an $r > 0$ such that $B_r(a) \cap E \subseteq U$.

ii) If $E$ is open, then $U$ is relatively open in $E$ if and only if $U$ is (plain old vanilla) open (in the usual sense).

**Proof.** i) If $U$ is relatively open in $E$, then $U = E \cap A$ for some open set $A$. Since $A$ is open, there is an $r > 0$ such that $B_r(a) \subseteq A$. Hence, $B_r(a) \cap E \subseteq A \cap E = U$.

Conversely, for each $a \in U$ choose an $r(a) > 0$ such that $B_{r(a)}(a) \cap E \subseteq U$. Then $\bigcup_{a \in U} B_{r(a)}(a) \cap E \subseteq U$. Since the union is taken over all $a \in U$, the reverse set inequality is also true. Thus $\bigcup_{a \in U} B_{r(a)}(a) \cap E = U$. Since the union of these open balls is open by Theorem 8.24, it follows that $U$ is relatively open in $E$. 
ii) Suppose that $U$ is relatively open in $E$. If $E$ and $A$ are open, then $U = E \cap A$ is open. Thus $U$ is open in the usual sense. Conversely, if $U$ is open, then $E \cap U = U$ is open. Thus every open subset of $E$ is relatively open in $E$. ■

Next, we introduce connectivity.

### 8.28 Definition.

Let $E$ be a subset of $\mathbb{R}^n$.

i) A pair of sets $U$, $V$ is said to separate $E$ if and only if $U$ and $V$ are nonempty, relatively open in $E$, $E = U \cup V$, and $U \cap V = \emptyset$.

ii) $E$ is said to be connected if and only if $E$ cannot be separated by any pair of relatively open sets $U$, $V$.

Loosely speaking, a connected set is all in one piece (i.e., cannot be broken into smaller, nonempty, relatively open pieces which do not share any common points).

The empty set is connected, since it can never be written as the union of nonempty sets. Every singleton $E = \{a\}$ is also connected, since if $E = U \cup V$, where $U \cap V = \emptyset$ and both $U$ and $V$ are nonempty, then $E$ has at least two points. More complicated connected sets can be found in the exercises.

Notice that by Definitions 8.26 and 8.28, a set $E$ is not connected if there are open sets $A$, $B$ such that $E \cap A$, $E \cap B$ are nonempty, $E = (E \cap A) \cup (E \cap B)$, and $A \cap B = \emptyset$. Is this statement valid if we replace $E = (E \cap A) \cup (E \cap B)$ by $E \subseteq A \cup B$?

### 8.29 Remark. Let $E \subseteq \mathbb{R}^n$. If there exists a pair of open sets $A, B$ such that $E \cap A \neq \emptyset$, $E \cap B \neq \emptyset$, $E \subseteq A \cup B$, and $A \cap B = \emptyset$, then $E$ is not connected.

**Proof.** Set $U = E \cap A$ and $V = E \cap B$. By hypothesis and Definition 8.26, $U$ and $V$ are relatively open in $E$ and nonempty. Since $U \cap V \subseteq A \cap B = \emptyset$, it suffices by Definition 8.28 to prove that $E = U \cup V$. But $E$ is a subset of $A \cup B$, so $E \subseteq U \cup V$. On the other hand, both $U$ and $V$ are subsets of $E$, so $E \supseteq U \cup V$. We conclude that $E = U \cup V$. ■

(The converse of this result is also true, but harder to prove—see Theorem 8.38 below.)

In practice, Remark 8.29 is often easier to apply than Definition 8.28. Here are several examples. The set $Q$ is not connected: set $A = (-\infty, \sqrt{2})$ and $B = (\sqrt{2}, \infty)$. The "bow-tie-shaped set" $\{(x, y) : -1 \leq x \leq 1 \text{ and } -|x| < y < |x|\}$ is not connected (see Figure 8.6): set $A = \{(x, y) : x < 0\}$ and $B = \{(x, y) : x > 0\}$.

Is there a simple description of all connected subsets of $\mathbb{R}^n$?
8.30 Theorem. A subset $E$ of $\mathbb{R}$ is connected if and only if $E$ is an interval.

Proof. Suppose that $E$ is a connected subset of $\mathbb{R}$. If $E$ is empty or if $E$ contains only one point $c$, then $E$ is one of the intervals $(c, c)$ or $[c, c]$.

Suppose that $E$ contains at least two points. Set $a = \inf E$ and $b = \sup E$, and observe that $-\infty \leq a < b \leq \infty$. If $a \in E$ set $a_k = a$, and if $b \in E$ set $b_k = b$, $k \in \mathbb{N}$. Otherwise, use the Approximation Property to choose $a_k, b_k \in E$ such that $a_k \downarrow a$ and $b_k \uparrow b$ as $k \to \infty$. Notice that in all cases, $E$ contains each $[a_k, b_k]$. Indeed, if not, say there is an $x \in [a_k, b_k] \setminus E$, then $a_k \in E \cap (-\infty, x)$, $b_k \in E \cap (x, \infty)$, and $E \subseteq (-\infty, x) \cup (x, \infty)$. Hence, by Remark 8.29, $E$ is not connected, a contradiction. Therefore, $E = [a_k, b_k]$ for all $k \in \mathbb{N}$. It follows from construction that

$$E = \bigcup_{k=1}^{\infty} [a_k, b_k].$$

Since this last union is either $(a, b)$, $[a, b]$, $(a, b]$, or $[a, b]$, we conclude that $E$ is an interval.

Conversely, suppose that $E$ is an interval which is not connected. Then there are sets $U, V$, relatively open in $E$, which separate $E$ (i.e., $E = U \cup V$, $U \cap V = \emptyset$), and there exist points $x_1 \in U$ and $x_2 \in V$. We may suppose that $x_1 < x_2$. Since $x_1, x_2 \in E$ and $E$ is an interval, $I_0 := [x_1, x_2] \subseteq E$. Define $f$ on $I_0$ by

$$f(x) = \begin{cases} 0 & x \in U \\ 1 & x \in V. \end{cases}$$

Since $U \cap V = \emptyset$, $f$ is well defined. We claim that $f$ is continuous on $I_0$. Indeed, fix $x_0 \in [x_1, x_2]$. Since $U \cup V = E \supseteq I_0$, it is evident that $x_0 \in U$.
or \( x_0 \in V \). We may suppose the former. Let \( y_k \in I_0 \) and suppose that \( y_k \to x_0 \) as \( k \to \infty \). Since \( U \) is relatively open, there is an \( \varepsilon > 0 \) such that \((x_0 - \varepsilon, x_0 + \varepsilon) \cap E \subseteq U \). Since \( y_k \in E \) and \( y_k \to x_0 \), it follows that \( y_k \in U \) for large \( k \).
Hence \( f(y_k) = 0 = f(x_0) \) for large \( k \). Therefore, \( f \) is continuous at \( x_0 \) by the Sequential Characterization of Continuity.
We have proved that \( f \) is continuous on \( I_0 \). Hence by the Intermediate Value Theorem (Theorem 3.29), \( f \) must take on the value \( 1/2 \) somewhere on \( I_0 \). This is a contradiction, since by construction, \( f \) takes on only the values 0 or 1.

We shall use this result later to prove that a real function is continuous on a closed, bounded interval if and only if its graph is closed and connected (see Theorem 9.51).

EXERCISES

8.3.1. Sketch each of the following sets. Identify which of the following sets are open, which are closed, and which are neither. Also discuss the connectivity of each set.

a) \( E = \{(x, y) : y \neq 0\} \)
b) \( E = \{(x, y) : x^2 + 4y^2 \leq 1\} \)
c) \( E = \{(x, y) : y \geq x^2, 0 \leq y < 1\} \)
d) \( E = \{(x, y) : x^2 - y^2 > 1, -1 < y < 1\} \)
e) \( E = \{(x, y) : x^2 - 2x + y^2 = 0\} \cup \{(x, 0) : x \in [2, 3]\} \)

8.3.2. Let \( n \in \mathbb{N} \), let \( a \in \mathbb{R}^n \), let \( s, r \in \mathbb{R} \) with \( s < r \), and set

\[ V = \{x \in \mathbb{R}^n : s < \|x - a\| < r\} \quad \text{and} \quad E = \{x \in \mathbb{R}^n : s \leq \|x - a\| \leq r\}. \]

Prove that \( V \) is open and \( E \) is closed.

8.3.3. a) Let \( a \leq b \) and \( c \leq d \) be real numbers. Sketch a graph of the rectangle

\[ [a, b] \times [c, d] := \{(x, y) : x \in [a, b], y \in [c, d]\}, \]

and decide whether this set is connected. Explain your answers.
b) Sketch a graph of set

\[ B_1(-2, 0) \cup B_1(2, 0) \cup \{(x, 0) : -1 < x < 1\}, \]

and decide whether this set is connected. Explain your answers.

8.3.4. a) Set \( E_1 := \{(x, y) : y \geq 0\} \) and \( E_2 := \{(x, y) : x^2 + 2y^2 < 6\}, \) and sketch a graph of the set

\[ U := \{(x, y) : x^2 + 2y^2 < 6 \quad \text{and} \quad y \geq 0\}. \]
b) Decide whether $U$ is relatively open or relatively closed in $E_1$. Explain your answer.

c) Decide whether $U$ is relatively open or relatively closed in $E_2$. Explain your answer.

8.3.5. a) Let $E_1$ denote the closed ball centered at $(0, 0)$ of radius 1 and $E_2 := B_{\sqrt{2}}(2, 0)$, and sketch a graph of the set

$$U := \{(x, y) : x^2 + y^2 \leq 1 \text{ and } x^2 - 4x + y^2 + 2 < 0\}.$$

b) Decide whether $U$ is relatively open or relatively closed in $E_1$. Explain your answer.

c) Decide whether $U$ is relatively open or relatively closed in $E_2$. Explain your answer.

8.3.6. Suppose that $E \subseteq \mathbb{R}^n$ and that $C$ is a subset of $E$.

a) Prove that if $E$ is closed, then $C$ is relatively closed in $E$ if and only if $C$ is (plain old vanilla) closed (in the usual sense).

b) Prove that $C$ is relatively closed in $E$ if and only if $E \setminus C$ is relatively open in $E$.

8.3.7. a) If $A$ and $B$ are connected in $\mathbb{R}^n$ and $A \cap B \neq \emptyset$, prove that $A \cup B$ is connected.

b) If $\{E_\alpha : \alpha \in A\}$ is a collection of connected sets in $\mathbb{R}^n$ and $\bigcap_{\alpha \in A} E_\alpha \neq \emptyset$, prove that

$$E = \bigcup_{\alpha \in A} E_\alpha$$

is connected.

c) If $A$ and $B$ are connected in $\mathbb{R}$ and $A \cap B \neq \emptyset$, prove that $A \cap B$ is connected.

d) Show that part c) is no longer true if $\mathbb{R}^2$ replaces $\mathbb{R}$.

8.3.8. Let $V$ be a subset of $\mathbb{R}^n$.

a) Prove that $V$ is open if and only if there is a collection of open balls $
\{B_\alpha : \alpha \in A\}$ such that

$$V = \bigcup_{\alpha \in A} B_\alpha.$$

b) What happens to this result when open is replaced by closed?

8.3.9. Show that if $E$ is closed in $\mathbb{R}^n$ and $a \notin E$, then

$$\inf_{x \in E} \|x - a\| > 0.$$

8.3.10. Graph generic open balls in $\mathbb{R}^2$ with respect to each of the "non-Euclidean" norms $\| \cdot \|_1$ and $\| \cdot \|_\infty$. What shape are they?
8.4 INTERIOR, CLOSURE, AND BOUNDARY

To prove that every set contains a largest open set and is contained in a smallest closed set, we introduce the following topological operations.

8.31 Definition.

Let $E$ be a subset of a Euclidean space $\mathbb{R}^n$.

i) The interior of $E$ is the set

$$E^o := \bigcup \{ V : V \subseteq E \text{ and } V \text{ is open in } \mathbb{R}^n \}.$$ 

ii) The closure of $E$ is the set

$$\overline{E} := \bigcap \{ B : B \supseteq E \text{ and } B \text{ is closed in } \mathbb{R}^n \}.$$ 

Notice that every set $E$ contains the open set $\emptyset$ and is contained in the closed set $\mathbb{R}^n$. Hence, the sets $E^o$ and $\overline{E}$ are well defined. Also notice by Theorem 8.24 that the interior of a set is always open and the closure of a set is always closed.

The following result shows that $E^o$ is the largest open set contained in $E$, and $\overline{E}$ is the smallest closed set which contains $E$.

8.32 Theorem. Let $E \subseteq \mathbb{R}^n$. Then

i) $E^o \subseteq E \subseteq \overline{E}$,

ii) if $V$ is open and $V \subseteq E$, then $V \subseteq E^o$, and

iii) if $C$ is closed and $C \supseteq E$, then $C \supseteq \overline{E}$.

Proof: Since every open set $V$ in the union defining $E^o$ is a subset of $E$, it is clear that the union of these $V$'s is a subset of $E$. Thus $E^o \subseteq E$. A similar argument establishes $E \subseteq \overline{E}$. This proves i).

By Definition 8.31, if $V$ is an open subset of $E$, then $V \subseteq E^o$, and if $C$ is a closed set containing $E$, then $E \subseteq C$. This proves ii) and iii).

In particular, the interior of a bounded interval with endpoints $a$ and $b$ is $(a, b)$, and its closure is $[a, b]$. In fact, it is evident by parts ii) and iii) that $E = E^o$ if and only if $E$ is open, and $E = \overline{E}$ if and only if $E$ is closed. We shall use this observation many times below.

Let us examine these concepts in the concrete setting $\mathbb{R}^2$.

8.33 EXAMPLES.

a) Find the interior and closure of the set $E = \{(x, y) : -1 \leq x \leq 1 \text{ and } -|x| < y < |x|\}$.

b) Find the interior and closure of the set $E = B_1(-2, 0) \cup B_1(2, 0) \cup \{(x, 0) : -1 \leq x \leq 1\}$. 

Solution.

a) Graph \( y = |x| \) and \( x = \pm 1 \), and observe that \( E \) is a bow tie-shaped region with "solid" vertical edges (see Figure 8.6). Now, by Definition 8.20, any open set in \( \mathbb{R}^2 \) must contain a disk around each of its points. Since \( E^o \) is the largest open set inside \( E \), it is clear that

\[
E^o = \{(x, y) : -1 < x < 1 \text{ and } -|x| < y < |x|\}.
\]

Similarly,

\[
\overline{E} = \{(x, y) : -1 \leq x \leq 1 \text{ and } -|x| \leq y \leq |x|\}.
\]

b) Draw a graph of this region. It turns out to be "dumbbell shaped"; two open disks joined by a straight line. Thus \( E^o = B_1(-2, 0) \cup B_1(2, 0) \), and

\[
\overline{E} = B_1(-2, 0) \cup B_1(2, 0) \cup \{(x', 0) : -1 \leq x' \leq 1\}.
\]

These examples illustrate the fact that the interior of a nice enough set \( E \) in \( \mathbb{R}^2 \) can be obtained by removing all its "edges," and the closure of \( E \) by adding all its "edges."

One of the most important results from Chapter 5 is the Fundamental Theorem of Calculus. It states that the behavior of a derivative \( f' \) on an interval \( [a, b] \), as measured by its integral, is determined by the values of \( f \) at the endpoints of \( [a, b] \). What shall we use for "endpoints" of an arbitrary set in \( \mathbb{R}^n \)? Notice that the endpoints \( a, b \) are the only points which lie near both \( [a, b] \) and the complement of \( [a, b] \). Using this as a cue, we introduce the following concept.

8.34 Definition.

Let \( E \subseteq \mathbb{R}^n \). The boundary of \( E \) is the set

\[
\partial E := \{x \in \mathbb{R}^n : \text{ for all } r > 0, \ B_r(x) \cap E \neq \emptyset \text{ and } B_r(x) \cap E^c \neq \emptyset\}.
\]

[We will refer to the last two conditions in the definition of \( \partial E \) by saying that \( B_r(x) \) intersects \( E \) and \( E^c \).]

8.35 Example.

Describe the boundary of the set

\[
E = \{(x, y) : x^2 + y^2 \leq 9 \text{ and } (x - 1)(y + 2) > 0\}.
\]

Solution. Graph the relations \( x^2 + y^2 = 9 \) and \( (x - 1)(y + 2) = 0 \) to see that \( E \) is a region with solid curved edges and dashed straight edges (see Figure 8.7). By definition, then, the boundary of \( E \) is the union of these curved and straight edges (all made solid). Rather than describing \( \partial E \) analytically (which would
Section 8.4  Interior, Closure, and Boundary 299

![Figure 8.7]

involve solving for the intersection points of the straight lines $x = 1$, $y = -2$, and the circle $x^2 + y^2 = 9$, it is easier to describe $\partial E$ by using set algebra.

$$
\partial E = \{(x, y) : x^2 + y^2 \leq 9 \text{ and } (x - 1)(y + 2) \geq 0\} \cup \{(x, y) : x^2 + y^2 < 9 \text{ and } (x - 1)(y + 2) > 0\}
$$

It turns out that set algebra can be used to describe the boundary of any set.

8.36 Theorem. Let $E \subseteq \mathbb{R}^n$. Then $\partial E = \overline{E} \setminus \text{int } E$.

Proof. By Definition 8.34, it suffices to show that

- $x \in \overline{E}$ if and only if $B_r(x) \cap E \neq \emptyset$ for all $r > 0$, and
- $x \notin \text{int } E$ if and only if $B_r(x) \cap E^c \neq \emptyset$ for all $r > 0$. (9)

We will provide the details for (9) and leave the proof of (10) as an exercise. Suppose that $x \in \overline{E}$ but that $B_{r_0}(x) \cap E = \emptyset$ for some $r_0 > 0$. Then $(B_{r_0}(x))^c$ is a closed set which contains $E$; hence, by Theorem 8.32(iii), $\overline{E} \subseteq (B_{r_0}(x))^c$. It follows that $\overline{E} \cap B_{r_0}(x) = \emptyset$ (e.g., $x \notin \overline{E}$), a contradiction.

Conversely, suppose that $x \notin \overline{E}$. Since $(\overline{E})^c$ is open, there is an $r_0 > 0$ such that $B_{r_0}(x) \subseteq (\overline{E})^c$. In particular, $\emptyset = B_{r_0}(x) \cap \overline{E} \supseteq B_{r_0}(x) \cap E$ for some $r_0 > 0$. (10)

We have introduced topological operations (interior, closure, and boundary). The following result answers the question, How do these operations interact with the set operations (union and intersection)?
8.37 Theorem. Let $A, B \subseteq \mathbb{R}^n$. Then

i) \[(A \cup B)^o \supseteq A^o \cup B^o, \quad (A \cap B)^o = A^o \cap B^o,\]

ii) \[\overline{A \cup B} = \overline{A} \cup \overline{B}, \quad \overline{A \cap B} \subseteq \overline{A} \cap \overline{B},\]

iii) \[\partial (A \cup B) \subseteq \partial A \cup \partial B, \quad \partial (A \cap B) \subseteq \partial A \cap \partial B.\]

Proof. i) Since the union of two open sets is open, $A^o \cup B^o$ is an open subset of $A \cup B$. Hence, by Theorem 8.32ii, $A^o \cup B^o \subseteq (A \cup B)^o$.

Similarly, $(A \cap B)^o \supseteq A^o \cap B^o$. On the other hand, if $V \subset A \cap B$, then $V \subset A$ and $V \subset B$. Thus $(A \cap B)^o \subseteq A^o \cap B^o$.

ii) Since $\overline{A} \cup \overline{B}$ is closed and contains $A \cup B$, it is clear that, by Theorem 8.32iii, $\overline{A \cup B} \subseteq \overline{A} \cup \overline{B}$. Similarly, $\overline{A \cap B} \subseteq \overline{A} \cap \overline{B}$. To prove the reverse inequality for union, suppose that $x \notin A \cup B$. Then, by Definition 8.31, there is a closed set $E$ which contains $A \cup B$ such that $x \notin E$. Since $E$ contains both $A$ and $B$, it follows that $x \notin \overline{A}$ and $x \notin \overline{B}$. This proves part ii).

iii) Let $x \in \partial (A \cup B)$; that is, suppose that $B_r(x)$ intersects $A \cup B$ and $(A \cup B)^c$ for all $r > 0$. Since $(A \cup B)^c = A^c \cap B^c$, it follows that $B_r(x)$ intersects both $A^c$ and $B^c$ for all $r > 0$. Thus $B_r(x)$ intersects $A$ and $A^c$ for all $r > 0$, or $B_r(x)$ intersects $B$ and $B^c$ for all $r > 0$ (i.e., $x \in \partial A \cup \partial B$). This proves the first set inequality in part iii). A similar argument establishes the second inequality in part iii).

The second inequality in part iii) can be improved (see Exercise 8.4.10d).

Finally, we note (Exercise 8.4.11) that relatively open sets in $E$ can be divided into two kinds: those inside $E^o$, which contain none of their boundary, and those which intersect $\partial E$, which contain only that part of their boundary which intersects $\partial E$. (See Figures 15.3 and 15.4 for an illustration of both types.)

We close this section by showing that the converse of Remark 8.30 is also true. This result is optional because we do not use it anywhere else.

*8.38 Theorem. Let $E \subseteq \mathbb{R}^n$. If there exist nonempty, relatively open sets $U, V$ which separate $E$, then there is a pair of open sets $A, B$ such that $A \cap E \neq \emptyset$, $B \cap E \neq \emptyset$, $A \cap B = \emptyset$, and $E \subseteq A \cup B$.

Proof. We first show that

\[U \cap V = \emptyset. \tag{11}\]

Indeed, since $V$ is relatively open in $E$, there is a set $\Omega$, open in $\mathbb{R}^n$, such that $V = E \cap \Omega$. Since $U \cap V = \emptyset$, it follows that $U \subseteq \Omega^c$. This last set is closed in $\mathbb{R}^n$. Therefore,

\[\overline{U} \subseteq \overline{\Omega^c} = \Omega^c;\]

that is, (11) holds.
Next, we use (11) to construct the open set $B$. Set
\[
\delta_x := \inf\{\|x - u\| : u \in \overline{U} \}, \quad x \in V, \quad \text{and} \quad B = \bigcup_{x \in V} B_{\delta_x/2}(x).
\]
Clearly, $B$ is open in $\mathbb{R}^n$. Since $\delta_x > 0$ for each $x \not\in \overline{U}$ (see Exercise 8.3.9), $B$ contains $V$; hence $B \cap E \subseteq V$. The reverse inequality also holds, since by construction $B \cap U = \emptyset$ and by hypothesis $E = U \cup V$. Therefore, $B \cap E = V$.

Similarly, we can construct an open set $A$ such that $A \cap E = U$ by setting
\[
\epsilon_y := \inf\{\|v - y\| : v \in \overline{V} \}, \quad y \in U \quad \text{and} \quad A = \bigcup_{y \in U} B_{\epsilon_y/2}(y).
\]
In particular, $A$ and $B$ are nonempty open sets which satisfy $E \subseteq A \cup B$.

It remains to prove that $A \cap B = \emptyset$. Suppose, to the contrary, that there is a point $a \in A \cap B$. Then $a \in B_{\delta_x/2}(x)$ for some $x \in V$ and $a \in B_{\epsilon_y/2}(y)$ for some $y \in U$. We may suppose that $\delta_x \leq \epsilon_y$. Then
\[
\|x - y\| \leq \|x - a\| + \|a - y\| < \frac{\delta_x}{2} + \frac{\epsilon_y}{2} \leq \epsilon_y.
\]
Therefore, $\|x - y\| < \inf\{\|v - y\| : v \in \overline{V} \}$. Since $x \in V$, this is impossible. We conclude that $A \cap B = \emptyset$.

**EXERCISES**

**8.4.1.** Find the interior, closure, and boundary of each of the following subsets of $\mathbb{R}$.

a) $E = \{1/n : n \in \mathbb{N}\}$

b) $E = \bigcup_{n=1}^{\infty} \left( -\frac{1}{n+1}, \frac{1}{n} \right)$

c) $E = \bigcup_{n=1}^{\infty} (-n, n)$

d) $E = \mathbb{Q}$

**8.4.2.** For each of the following sets, sketch $E^o$, $\overline{E}$, and $\partial E$.

a) $E = \{(x, y) : x^2 + 4y^2 \leq 1\}$

b) $E = \{(x, y) : x^2 - 2x + y^2 = 0\} \cup \{(x, 0) : x \in [2, 3]\}$

c) $E = \{(x, y) : y \geq x^2, 0 \leq y < 1\}$

d) $E = \{(x, y) : x^2 - y^2 < 1, -1 < y < 1\}$

**8.4.3.** This exercise is used in Section 12.1. Suppose that $A \subseteq B \subseteq \mathbb{R}^n$.

Prove that
\[
\overline{A} \subseteq \overline{B} \quad \text{and} \quad A^o \subseteq B^o.
\]

**8.4.4.** Let $E$ be a subset of $\mathbb{R}^n$.

a) Prove that every subset $A \subseteq E$ contains a set $B$ which is the largest subset of $A$ that is relatively open in $E$. 
Chapter 8. Euclidean Spaces

b) Prove that every subset $A \subseteq E$ is contained in a set $B$ which is the smallest closed set containing $A$ that is relatively closed in $E$.

8.4.5. Complete the proof of Theorem 8.36 by verifying (10).

8.4.6. Prove that if $E \subseteq \mathbb{R}$ is connected, then $E^o$ is also connected. Show that this is false if "$\mathbb{R}$" is replaced by "$\mathbb{R}^2$".

8.4.7. Suppose that $E \subseteq \mathbb{R}^n$ is connected and that $E \subseteq A \subseteq \bar{E}$. Prove that $A$ is connected.

8.4.8. A set $A$ is called clopen if and only if it is both open and closed.

a) Prove that every Euclidean space has at least two clopen sets.

b) Prove that a proper subset $E$ of $\mathbb{R}^n$ is connected if and only if it contains exactly two relatively clopen sets.

c) Prove that every nonempty proper subset of $\mathbb{R}^n$ has a nonempty boundary.

8.4.9. Show that Theorem 8.37 is best possible in the following sense.

a) There exist sets $A, B$ in $\mathbb{R}$ such that $(A \cup B)^o \neq A^o \cup B^o$.

b) There exist sets $A, B$ in $\mathbb{R}$ such that $A \cap B \neq \bar{A} \cap \bar{B}$.

b) There exist sets $A, B$ in $\mathbb{R}$ such that $\partial(A \cup B) \neq \partial A \cup \partial B$ and $\partial(A \cap B) \neq \partial A \cup \partial B$.

8.4.10. Let $A$ and $B$ be subsets of $\mathbb{R}^n$.

a) Show that $\partial(A \cap B) \cap (A^c \cup (B^c)^*) \subseteq \partial A$.

b) Show that if $x \in \partial(A \cap B)$ and $x \notin (A \cap B) \cup (B \cap A)$, then $x \in \partial A \cap \partial B$.

c) Prove that $\partial(A \cap B) \subseteq (A \cap \partial B) \cup (B \cap \partial A) \cup (\partial A \cap \partial B)$.

d) Show that even in $\mathbb{R}$, there exist sets $A$ and $B$ such that $\partial(A \cap B) \neq (A \cap \partial B) \cup (B \cap \partial A) \cup (\partial A \cap \partial B)$.

8.4.11. Let $E \subseteq \mathbb{R}^n$ and $U$ be relatively open in $E$.

a) If $U \subseteq E^o$, then $U \cap \partial U = \emptyset$.

b) If $U \cap \partial E \neq \emptyset$, then $U \cap \partial U = U \cap \partial E$. 