8

DIFFERENTIAL CALCULUS OF SCALAR AND VECTOR FIELDS

8.1 Functions from $\mathbf{R}^n$ to $\mathbf{R}^m$. Scalar and vector fields

Part 1 of this volume dealt primarily with linear transformations

$$T: V \to W$$

from one linear space V into another linear space W. In Part 2 we drop the requirement that T be linear but restrict the spaces V and W to be finite-dimensional. Specifically, we shall consider functions with domain in n-space $\mathbf{R}^n$ and with range in m-space $\mathbf{R}^m$.

When both n and m are equal to 1, such a function is called a real-valued function of a real variable. When $n = 1$ and $m > 1$ it is called a vector-valued function of a real variable. Examples of such functions were studied extensively in Volume I.

In this chapter we assume that $n > 1$ and $m \geq 1$. When $m = 1$ the function is called a real-valued function of a vector variable or, more briefly, a *scalar field.* When $m > 1$ it is called a vector-valued function of a vector variable, or simply a *vector field.*

This chapter extends the concepts of limit, continuity, and derivative to scalar and vector fields. Chapters 10 and 11 extend the concept of the integral.

Notation: Scalars will be denoted by light-faced type, and vectors by bold-faced type. If f is a scalar field defined at a point $\mathbf{x} = (x_1, \ldots, x_n)$ in $\mathbf{R}^n$, the notations $f(\mathbf{x})$ and $f(x_1, \ldots, x_n)$ are both used to denote the value of f at that particular point. If $\mathbf{f}$ is a vector field we write $\mathbf{f}(\mathbf{x})$ or $\mathbf{f}(x_1, \ldots, x_n)$ for the function value at $\mathbf{x}$. We shall use the inner product

$$\mathbf{x} \cdot \mathbf{y} = \sum_{k=1}^{n} x_k y_k$$

and the corresponding norm $\|\mathbf{x}\| = (\mathbf{x} \cdot \mathbf{x})^{1/2}$, where $\mathbf{x} = (x_1, \ldots, x_n)$ and $\mathbf{y} = (y_1, \ldots, y_n)$. Points in $\mathbf{R}^2$ are usually denoted by (x, y) instead of (x_1, x_2); points in $\mathbf{R}^3$ by (x, y, z) instead of (x_1, x_2, x_3).

Scalar and vector fields defined on subsets of $\mathbf{R}^2$ and $\mathbf{R}^3$ occur frequently in the applications of mathematics to science and engineering. For example, if at each point $\mathbf{x}$ of the atmosphere we assign a real number $f(\mathbf{x})$ which represents the temperature at $\mathbf{x}$, the function

f so defined is a scalar field. If we assign a vector which represents the wind velocity at that point, we obtain an example of a vector field.

In physical problems dealing with either scalar or vector fields it is important to know how the field changes as we move from one point to another. In the one-dimensional case the derivative is the mathematical tool used to study such changes. Derivative theory in the one-dimensional case deals with functions defined on open intervals. To extend the theory to $\mathbf{R}^n$ we consider generalizations of open intervals called *open sets*.

8.2 Open balls and open sets

Let a be a given point in $\mathbf{R}^n$ and let r be a given positive number. The set of all points x in $\mathbf{R}^n$ such that

$$\|x - a\| < r$$

is called an *open n-ball* of radius r and center a. We denote this set by $B(a)$ or by $B(a; r)$.

The ball $B(a; r)$ consists of all points whose distance from a is less than r. In $\mathbf{R}^1$ this is simply an open interval with center at a. In $\mathbf{R}^2$ it is a circular disk, and in $\mathbf{R}^3$ it is a spherical solid with center at a and radius r.

DEFINITION OF AN INTERIOR POINT. *Let S be a subset of $\mathbf{R}^n$, and assume that $a \in S$. Then a is called an interior point of S if there is an open n-ball with center at a, all of whose points belong to S.*

In other words, every interior point a of S can be surrounded by an n-ball $B(a)$ such that $B(a) \subseteq S$. The set of all interior points of S is called the *interior* of S and is denoted by int S. An open set containing a point a is sometimes called a *neighborhood of a.*

DEFINITION OF AN OPEN SET. *A set S in $\mathbf{R}^n$ is called open if all its points are interior points. In other words, S is open if and only if $S = \text{int } S$.*

EXAMPLES. In $\mathbf{R}^1$ the simplest type of open set is an open interval. The union of two or more open intervals is also open. A closed interval $[a, b]$ is not an open set because neither endpoint of the interval can be enclosed in a 1-ball lying within the given interval.

The 2-ball $S = B(O; 1)$ shown in Figure 8.1 is an example of an open set in $\mathbf{R}^2$. Every point a of S is the center of a disk lying entirely in S. For some points the radius of this disk is very small.

Some open sets in $\mathbf{R}^2$ can be constructed by taking the Cartesian product of open sets in $\mathbf{R}^1$. If A_1 and A_2 are subsets of $\mathbf{R}^1$, their Cartesian product $A_1 \times A_2$ is the set in $\mathbf{R}^2$ defined by

$$A_1 \times A_2 = \{(a_1, a_2) \mid a_1 \in A_1 \text{ and } a_2 \in A_2\}.$$

An example is shown in Figure 8.2. The sets A_1 and A_2 are intervals, and $A_1 \times A_2$ is a rectangle.

If A_1 and A_2 are open subsets of $\mathbf{R}^1$, then $A_1 \times A_2$ will be an open subset of $\mathbf{R}^2$. To prove this, choose any point $a = (a_1, a_2)$ in $A_1 \times A_2$. We must show that a is an interior point

FIGURE 8.1 The disk $B(O; 1)$ is an open set in $\mathbf{R}^2$.

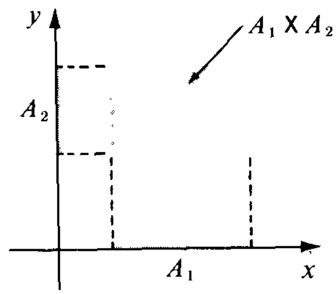

FIGURE 8.2 The Cartesian product of two open intervals is an open rectangle.

of A, x A,. Since A_1 and A_2 are open in $\mathbf{R}^1$ there is a l-ball $B(a_1; r_1)$ in A_1 and a l-ball $B(a_2 ; r_2)$ in A,. Let $r = \min \{r_1, r_2\}$. We can easily show that the 2-ball $B(\boldsymbol{a}; r) \subseteq A_1$ x A,. In fact, if $\boldsymbol{x} = (x_1, x_2)$ is any point of $B(\boldsymbol{a}; r)$ then $\|\boldsymbol{x} - \boldsymbol{a}\| < r$, so $|x_1 - a_1| < r_1$ and $|x_2 - a_2| < r_2$. Hence $x_1 \in B(a_1 ; r_1)$ and $x_2 \in B(a_2 ; r_2)$. Therefore $x_1 \in$ A, and $x_2 \in A_2$, so $(x_1, x_2) \in A_1$ x A,. This proves that every point of $B(\boldsymbol{a}; r)$ is in A, x A_2. Therefore every point of A, x A_2 is an interior point, so A, x A_2 is open.

The reader should realize that an open subset of $\mathbf{R}^1$ is no longer an open set when it is considered as a subset of $\mathbf{R}^2$, because a subset of $\mathbf{R}^1$ cannot contain a 2-ball.

DEFINITIONS OF EXTERIOR AND BOUNDARY. *A point $\boldsymbol{x}$ is said to be exterior to a set S in $\mathbf{R}^n$ if there is an n-ball B(x) containing no points of S. The set of all points in $\mathbf{R}^n$ exterior to S is culled the exterior of S and is denoted by* ext S. *A point which is neither exterior to S nor an interior point of S is called a boundary point of S. The set of all boundary points of S is called the boundary of S and is denoted by ∂S.*

These concepts are illustrated in Figure 8.1. The exterior of S is the set of all $\boldsymbol{x}$ with $\|\boldsymbol{x}\| > 1$. The boundary of S consists of all $\boldsymbol{x}$ with $\|\boldsymbol{x}\| = 1$.

8.3 Exercises

1. Letf be a scalar field defined on a set S and let c be a given real number. The set of all points x in S such that $f(\boldsymbol{x}) = c$ is called a *level set* off. (Geometric and physical problems dealing with level sets will be discussed later in this chapter.) For each of the following scalar fields. S is the whole space $\mathbf{R}^n$. Make a sketch to describe the level sets corresponding to the given values of c.
   (a) $f(x, y) = x^2 + y^2$, $\quad c = 0, 1, 4, 9$.
   (b) $f(x, y) = e^{xy}$, $\quad c = e^{-2}, e^{-1}, 1, e, e^2, e^3$.
   (c) $f(x, y) = \cos (x + y)$, $\quad c = -1, 0, \frac{1}{2}, \frac{1}{2}\sqrt{2}, 1$.
   (d) $f(x, y, z) = x + y + z$, $\quad c = -1, 0, 1$.
   (e) $f(x, y, z) = x^2 + 2y^2 + 3z^2$, $\quad c = 0, 6, 12$.
   (f) $f(x, y, z) = \sin (x^2 + y^2 + z^2)$, $\quad c = -1, -\frac{1}{2}, 0, \frac{1}{2}\sqrt{2}, 1$.

2. In each of the following cases, let S be the set of all points (x, y) in the plane satisfying the given inequalities. Make a sketch showing the set S and explain, by a geometric argument, whether or not S is open. Indicate the boundary of S on your sketch.
(a) $x^2 + y^2 < 1$.
(b) $3x^2 + 2y^2 < 6$.
(c) $|x| < 1$ and $|y| < 1$.
(d) $x \geq 0$ and $y > 0$.
(e) $|x| \leq 1$ and $|y| \leq 1$.
(f) $x > 0$ and $y < 0$.
(g) $xy < 1$.
(h) $1 \leq x \leq 2$ and $3 < y < 4$.
(i) $1 < x < 2$ and $y > 0$.
(j) $x \geq y$.
(k) $x > y$.
(1) $y > x^2$ and $|x| < 2$.
(m) $(x^2 + y^2 - 1)(4 - x^2 - y^2) > 0$.
(n) $(2x - x^2 - y^2)(x^2 + y^2 - x) > 0$.
3. In each of the following, let S be the set of all points (x, y, z) in 3-space satisfying the given inequalities and determine whether or not S is open.
(a) $z^2 - x^2 - y^2 - 1 > 0$.
(b) $|x| < 1$, $|y| < 1$, and $|z| < 1$.
(c) $x + y + z < 1$.
(d) $|x| \leq 1$, $|y| < 1$, and $|z| < 1$.
(e) $x + y + z < 1$ and $x > 0, y > 0, z > 0$.
(f) $x^2 + 4y^2 + 4z^2 - 2x + 16y + 40z + 113 < 0$.
4. (a) If A is an open set in n-space and if $x \in A$, show that the set $A - \{x\}$, obtained by removing the point x from A, is open.
(b) If A is an open interval on the real line and B is a closed subinterval of A, show that $A - B$ is open.†
(c) If A and B are open intervals on the real line, show that $A \cup B$ and $A \cap B$ are open.
(d) If A is a closed interval on the real line, show that its complement (relative to the whole real line) is open.
5. Prove the following properties of open sets in $\mathbf{R}^n$:
(a) The empty set $\varnothing$ is open.
(b) $\mathbf{R}^n$ is open.
(c) The union of any collection of open sets is open.
(d) The intersection of a finite collection of open sets is open.
(e) Give an example to show that the intersection of an infinite collection of open sets is not necessarily open.

Closed sets. A set Sin $\mathbf{R}^n$ is called *closed* if its complement $\mathbf{R}^n - S$ is open. The next three exercises discuss properties of closed sets.

6. In each of the following cases, let S be the set of all points (x, y) in $\mathbf{R}^2$ satisfying the given conditions. Make a sketch showing the set Sand give a geometric argument to explain whether S is open, closed, both open and closed, or neither open nor closed.
(a) $x^2 + y^2 \geq 0$.
(b) $x^2 + y^2 < 0$.
(c) $x^2 + y^2 \leq 1$.
(d) $1 < x^2 + y^2 < 2$.
(e) $1 \leq x^2 + y^2 \leq 2$.
(f) $1 < x^2 + y^2 \leq 2$.
(g) $1 \leq x \leq 2, 3 \leq y \leq 4$.
(h) $1 \leq x \leq 2, 3 \leq y < 4$.
(i) $y = x^2$.
(j) $y \geq x^2$.
(k) $y \geq x^2$ and $|x| < 2$.
(1) $y \geq x^2$ and $|x| \leq 2$.
7. (a) If A is a closed set in n-space and x is a point not in A, prove that $A \cup \{x\}$ is also closed.
(b) Prove that a closed interval $[a, b]$ on the real line is a closed set.
(c) If A and B are closed intervals on the real line, show that $A \cup B$ and $A \cap B$ are closed.

† If A and B are sets, the difference $A - B$ (called the ***complement of B relative to A)*** is the set of all elements of A which are not in B.

8. Prove the following properties of closed sets in $\mathbf{R}^n$. You may use the results of Exercise 5.
   (a) The empty set $\varnothing$ is closed.
   (b) $\mathbf{R}^n$ is closed.
   (c) The intersection of any collection of closed sets is closed.
   (d) The union of a finite number of closed sets is closed.
   (e) Give an example to show that the union of an infinite collection of closed sets is not necessarily closed.
9. Let S be a subset of $\mathbf{R}^n$.
   (a) Prove that both int S and ext S are open sets.
   (b) Prove that $\mathbf{R}^n = (\text{int } S) \cup (\text{ext } S) \cup \partial S$, a union of disjoint sets, and use this to deduce that the boundary ∂S is always a closed set.
10. Given a set S in $\mathbf{R}^n$ and a point x with the property that every ball $B(x)$ contains both interior points of S and points exterior to S. Prove that x is a boundary point of S. Is the converse statement true? That is, does every boundary point of S necessarily have this property?
11. Let S be a subset of $\mathbf{R}^n$. Prove that ext $S = \text{int}(\mathbf{R}^n - S)$.
12. Prove that a set S in $\mathbf{R}^n$ is closed if and only if $S = (\text{int } S) \cup \partial S$.

8.4 Limits and continuity

The concepts of limit and continuity are easily extended to scalar and vector fields. We shall formulate the definitions for vector fields; they apply also to scalar fields.

We consider a function $f: S \to \mathbf{R}^m$, where S is a subset of $\mathbf{R}^n$. If $a \in \mathbf{R}^n$ and $b \in \mathbf{R}^m$ we write

$$(8.1) \qquad \lim_{x \to a} f(x) = b \qquad (\text{or}, f(x) \to b \text{ as } x \to a)$$

to mean that

$$(8.2) \qquad \lim_{\|x-a\| \to 0} \| f(x) - b\| = 0.$$

The limit symbol in equation (8.2) is the usual limit of elementary calculus. In this definition it is not required that f be defined at the point a itself.

If we write $h = x - a$, Equation (8.2) becomes

$$\lim_{\|h\| \to 0} \|f(a + h) - b\| = 0.$$

For points in $\mathbf{R}^2$ we write (x, y) for x and (a, b) for a and express the limit relation (8.1) as follows :

$$\lim_{(x,y) \to (a,b)} f(x, y) = b.$$

For points in $\mathbf{R}^3$ we put $x = (x, y, z)$ and $a = (a, b, c)$ and write

$$\lim_{(x,y,z) \to (a,b,c)} f(x, y, z) = b .$$

A function f is said to be *continuous* at a if f is defined at a and if

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

We say f is *continuous on a set S* if f is continuous at each point of S.

Since these definitions are straightforward extensions of those in the one-dimensional case, it is not surprising to learn that many familiar properties of limits and continuity can also be extended. For example, the usual theorems concerning limits and continuity of sums, products, and quotients also hold for scalar fields. For vector fields, quotients are not defined but we have the following theorem concerning sums, multiplication by scalars, inner products, and norms.

THEOREM 8.1. *If* $\lim_{x \to a} f(x) = b$ *and* $\lim_{x \to a} g(x) = c$, *then we also have:*

(a) $\lim_{x \to a} [f(x) + g(x)] = b + c$.

(b) $\lim_{x \to a} \lambda f(x) = \lambda b$ *for every scalar* λ.

(c) $\lim_{x \to a} f(x) \cdot g(x) = b \cdot c$.

(d) $\lim_{x \to a} \|f(x)\| = \|b\|$.

Proof. We prove only parts (c) and (d); proofs of (a) and (b) are left as exercises for the reader.

To prove (c) we write

$$f(x) \cdot g(x) - b \cdot c = [f(x) - b] \cdot [g(x) - c] + b \cdot [g(x) - c] + c \cdot [f(x) - b].$$

Now we use the triangle inequality and the Cauchy-Schwarz inequality to obtain

$$0 \le \|f(x) \cdot g(x) - b \cdot c\| \le \|f(x) - b\| \, \|g(x) - c\| + \|b\| \, \|g(x) - c\| + \|c\| \, \|f(x) - b\| .$$

Since $\|f(x) - b\| \to 0$ and $\|g(x) - c\| \to 0$ as $x \to a$, this shows that $\|f(x) \cdot g(x) - b \cdot c\| \to 0$ as $x \to a$, which proves (c).

Taking $f(x) = g(x)$ in part (c) we find

$$\lim_{x \to a} \|f(x)\|^2 = \|b\|^2 ,$$

from which we obtain (d).

EXAMPLE 1. *Continuity and components of a vector field.* If a vector field f has values in $\mathbf{R}^m$, each function value $f(x)$ has m components and we can write

$$f(x) = (f_1(x), \ldots, f_m(x)) .$$

The m scalar fields $f_1, \ldots, f_m$ are called ***components*** of the vector field f. We shall prove that f is continuous at a point if, and only if, each component f_k is continuous at that point.

Let e_k denote the kth unit coordinate vector (all the components of e_k are 0 except the kth, which is equal to 1). Then $f_k(x)$ is given by the dot product

$$f_k(x) = f(x) \cdot e_k .$$

Therefore, part (c) of Theorem 8.1 shows that each point of continuity of f is also a point of continuity of f_k. Moreover, since we have

$$f(x) = \sum_{k=1}^{m} f_k(x)e_k ,$$

repeated application of parts (a) and (b) of Theorem 8.1 shows that a point of continuity of all m components $f_1, \ldots, f_m$ is also a point of continuity off.

EXAMPLE 2. *Continuity of the identity function.* The identity function, $f(x) = x$, is continuous everywhere in R". Therefore its components are also continuous everywhere in $\mathbf{R}^n$. These are the n scalar fields given by

$$f_1(x) = x_1, f_2(x) = x_2, \ldots, f_n(x) = x_n .$$

EXAMPLE 3. *Continuity of linear transformations.* Let $f: \mathbf{R}^n \to \mathbf{R}^m$ be a linear transformation. We will prove that f is continuous at each point a in R". By linearity we have

$$f(a + h) = f(a) + f(h) .$$

Therefore, it suffices to prove that $f(h) \to 0$ as $h \to 0$. Writing h in terms of its components we have $h = h_1e_1 + \cdots + h_ne_n$. Using linearity again we find $f(h) = h_1f(e_1) + \cdots + h_nf(e_n)$. This shows that $f(h) \to 0$ as $h \to 0$.

EXAMPLE 4. *Continuity of polynomials in n variables.* A scalar field P defined on $\mathbf{R}^n$ by a formula of the form

$$P(x) = \sum_{k_1=0}^{p_1} \cdots \sum_{k_n=0}^{p_n} c_{k_1 \cdots k_n} x_1^{k_1} \cdots x_n^{k_n}$$

is called a polynomial in n variables $x_1, \ldots, x_n$. A polynomial is continuous everywhere in $\mathbf{R}^n$ because it is a finite sum of products of scalar fields continuous everywhere in $\mathbf{R}^n$. For example, a polynomial in two variables x and y, given by

$$P(x, y) = \sum_{i=0}^{p} \sum_{j=0}^{q} c_{ij} x^i y^j$$

is continuous at every point (x, y) in $\mathbf{R}^2$.

EXAMPLE 5. *Continuity of rational functions.* A scalar field f given by $f(x) = P(x)/Q(x)$, where P and Q are polynomials in the components of x, is called a rational function. A rational function is continuous at each point where $Q(x) \neq 0$.

Further examples of continuous function can be constructed with the help of the next theorem, which is concerned with continuity of composite functions.

THEOREM 8.2. *Let f and g be functions such that the composite function $f \circ g$ is defined at a, where*

$$(f \circ g)(x) = f[g(x)] .$$

If g is continuous at a and if f is continuous at $g(a)$, then the composition $f \circ g$ is continuous at a.

Proof. Let $y = f(x)$ and 6 $= g(a)$. Then we have

$$f[g(x)] - f[g(a)] = f(y) - f(b) .$$

By hypothesis, $y \to b$ as $x \to a$, so we have

$$\lim_{\|x-a\| \to 0} \|f[g(x)] - f[g(a)]\| = \lim_{\|y-b\| \to 0} \|f(y) - f(b)\| = 0.$$

Therefore $\lim_{x \to a} f[g(x)] = f[g(a)]$, so $f \circ g$ is continuous at a.

EXAMPLE 6. The foregoing theorem implies continuity of scalar fields h, where $h(x, y)$ is given by formulas such as

$$\sin(x^2 y), \qquad \log(x^2 + y^2), \qquad \frac{x^{x+y}}{x+y}, \qquad \log[\cos(x^2 + y^2)] .$$

These examples are continuous at all points at which the functions are defined. The first is continuous at all points in the plane, and the second at all points except the origin. The third is continuous at all points (x, y) at which $x + y \neq 0$, and the fourth at all points at which $x^2 + y^2$ is not an odd multiple of $\pi/2$. [The set of (x, y) such that $x^2 + y^2 = n\pi/2$, $n = 1, 3, 5, \ldots$, is a family of circles centered at the origin.] These examples show that the discontinuities of a function of two variables may consist of isolated points, entire curves, or families of curves.

EXAMPLE 7. A function of two variables may be continuous in each variable separately and yet be discontinuous as a function of the two variables together. This is illustrated by the following example:

$$f(x, y) = \frac{xy}{x^2 + y^2} \quad \text{if } (x, y) \neq (0, 0), \; f(0, 0) = 0 .$$

For points (x, y) on the x-axis we have $y = 0$ and $f(x, y) = f(x, 0) = 0$, so the function has the constant value 0 everywhere on the x-axis. Therefore, if we put $y = 0$ and think of f as a function of x alone, f is continuous at $x = 0$. Similarly, f has the constant value 0 at all points on the y-axis, so if we put $x = 0$ and think off as a function of y alone, f is continuous at $y = 0$. However, as a function of two variables, f is not continuous at the origin. In fact, at each point of the line $y = x$ (except at the origin) the function has the constant value $\frac{1}{2}$ because $f(x, x) = x^2/(2x^2) = \frac{1}{2}$; since there are points on this line arbitrarily close to the origin and since $f(0, 0) \neq \frac{1}{2}$, the function is not continuous at (0, 0).

8.5 Exercises

The exercises in this section are concerned with limits and continuity of scalar fields defined on subsets of the plane.

1. In each of the following examples a scalar fieldfis defined by the given equation for all points (x, y) in the plane for which the expression on the right is defined. In each example determine the set of points (x, y) at whichfis continuous.

(a) $f(x, y) = x^4 + y^4 - 4x^2y^2$.

(b) $f(x, y) = \log (x^2 + y^2)$.

(c) $f(x, y) = \dfrac{1}{y} \cos x^2$.

(d) $f(x, y) = \tan \dfrac{x^2}{y}$.

(e) $f(x, y) = \arctan \dfrac{y}{x}$.

(f) $f(x, y) = \arcsin \dfrac{x}{\sqrt{x^2 + y^2}}$.

(g) $f(x, y) = \arctan \dfrac{x + y}{1 - xy}$.

(h) $f(x, y) = \dfrac{x}{\sqrt{x^2 + y^2}}$.

(i) $f(x, y) = x^{(y^2)}$.

(j) $f(x, y) = \arccos \sqrt{x/y}$.

2. If $\lim_{(x,y)\to(a,b)} f(x, y) = L$, and if the one-dimensional limits

$$\lim_{x\to a} f(x, y) \qquad \text{and} \qquad \lim_{y\to b} f(x, y)$$

both exist, prove that

$$\lim_{x\to a} [\lim_{y\to b} f(x, y)] = \lim_{y\to b} [\lim_{x\to a} f(x, y)] = L.$$

The two limits in this equation are called *iterated* limits; the exercise shows that the existence of the two-dimensional limit and of the two one-dimensional limits implies the existence and equality of the two iterated limits. (The converse is not always true. A counter example is given in Exercise 4.)

3. Let $f(x, y) = (x - y)/(x + y)$ if $x + y \neq 0$. Show that

$$\lim_{\text{e-o}} [\lim_{y\to 0} f(x, y)] = 1 \qquad \text{but that} \quad \lim_{\text{u-o}} [\lim_{x\to 0} f(x, y)] = -1.$$

Use this result with Exercise 2 to deduce that $f(x, y)$ does not tend to a limit as $(x, y) \to (0, 0)$.

4. Let

$$f(x, y) = \frac{x^2y^2}{x^2y^2 + (x - y)^2} \qquad \text{whenever } x^2y^2 + (x - y)^2 \neq 0.$$

Show that

$$\lim_{x\to 0} [\lim_{y\to 0} f(x, y)] = \lim_{y\to 0} [\lim_{x\to 0} f(x, y)] = 0$$

but that $f(x, y)$ does not tend to a limit as $(x, y) \to (0, 0)$. *[Hint:* Examinefon the line $y = x$.] This example shows that the converse of Exercise 2 is not always true.

5. Let

$$f(x, y) = \begin{cases} x \sin \dfrac{1}{y} & \text{if } y \neq 0, \\ 0 & \text{if } y = 0. \end{cases}$$

Show that $f(x, y) \to 0$ as $(x, y) \to (0, 0)$ but that

$$\lim_{y\to 0} [\lim_{x\to 0} f(x, y)] \neq \lim_{x\to 0} [\lim_{y\to 0} f(x, y)] .$$

Explain why this does not contradict Exercise 2.

6. If $(x, y) \neq (0, 0)$, let $f(x, y) = (x^2 - y^2)/(x^2 + y^2)$. Find the limit of $f(x, y)$ as $(x, y) \to (0, 0)$ along the line $y = mx$. Is it possible to define $f(0, 0)$ so as to makefcontinuous at (0, O)?
7. Let $f(x, y) = 0$ if $y \leq 0$ or if $y \geq x^2$ and let $f(x, y) = 1$ if $0 < y < x^2$. Show that $f(x, y) \to 0$ as $(x, y) \to (0, 0)$ along any straight line through the origin. Find a curve through the origin along which (except at the origin) $f(x, y)$ has the constant-value 1. Isfcontinuous at the origin?
8. If $f(x, y) = [\sin (x^2 + y^2)]/(x^2 + y^2)$ when $(x, y) \neq (0, 0)$ how must $f(0, 0)$ be defined so as to makefcontinuous at the origin?
9. Letfbe a scalar field continuous at an interior point a of a set S in R”. If $f(a) \neq 0$, prove that there is an n-ball $B(a)$ in whichfhas the same sign as $f(a)$.

8.6 The derivative of a scalar field with respect to a vector

This section introduces derivatives of scalar fields. Derivatives of vector fields are discussed in Section 8.18.

Let f be a scalar field defined on a set S in R”, and let a be an interior point of S. We wish to study how the field changes as we move from a to a nearby point. For example, suppose $f(a)$ is the temperature at a point a in a heated room with an open window. If we move toward the window the temperature will decrease, but if we move toward the heater it will increase. In general, the manner in which a field changes will depend on the direction in which we move away from a.

Suppose we specify this direction by another vector y. That is, suppose we move from a toward another point $a + y$ along the line segment joining a and $a + y$. Each point on this segment has the form $a + hy$, where h is a real number. An example is shown in Figure 8.3. The distance from a to $a + hy$ is $\|hy\| = |h| \, \|y\|$.

Since a is an interior point of S, there is an n-ball $B(a; r)$ lying entirely in S. If h is chosen so that $|h| \, \|y\| < r$, the segment from a to $a + hy$ will lie in S. (See Figure 8.4.) We keep

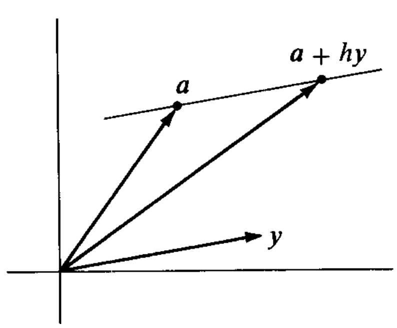

FIGURE 8.3 The point $a + hy$ lies on the line through a parallel to y.

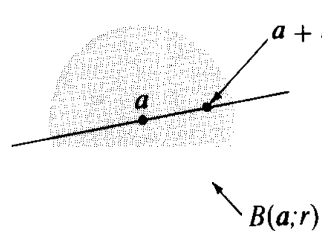

FIGURE 8.4 The point $a + hy$ lies in the n-ball $B(a; r)$ if $\|hy\| < r$.

$h \neq 0$ but small enough to guarantee that $a + hy \in S$ and we form the difference quotient

$$\frac{f(a + hy) - f(a)}{h}. \tag{8.3}$$

The numerator of this quotient tells us how much the function changes when we move from a to $a + hy$. The quotient itself is called the *average rate of change* off over the line segment joining a to $a + hy$. We are interested in the behavior of this quotient as $h \to 0$. This leads us to the following definition.

DEFINITION OF THE DERIVATIVE OF A SCALAR FIELD WITH RESPECT TO A VECTOR. *Given a scalar field $f: S \to R$, where $S \subseteq \mathbf{R}^n$. Let a be an interior point of S and let y be an arbitrary point in R^n. The derivative off at a with respect toy is denoted by the symbol $f'(a; y)$ and is defined by the equation*

$$f'(a; y) = \lim_{h \to 0} \frac{f(a + hy) - f(a)}{h} \tag{8.4}$$

when the limit on the right exists.

EXAMPLE 1. If $y = O$, the difference quotient (8.3) is 0 for every $h \neq 0$, so $f'(a; 0)$ always exists and equals 0.

EXAMPLE 2. *Derivative of a linear transformation.* If $f: S \to R$ is linear, then $f(a + hy) = f(a) + hf(y)$ and the difference quotient (8.3) is equal to $f(y)$ for every $h \neq 0$. In this case, $f'(a; y)$ always exists and is given by

$$f'(a; y) = f(y)$$

for every a in S and every y in R^n. In other words, the derivative of a linear transformation with respect to y is equal to the value of the function at y.

To study how f behaves on the line passing through a and $a + y$ for $y \neq O$ we introduce the function

$$g(t) = f(a + ty).$$

The next theorem relates the derivatives $g'(t)$ and $f'(a + ty; y)$.

THEOREM 8.3. *Let $g(t) = f(a + ty)$. If one of the derivatives $g'(t)$ or $f'(a + ty; y)$ exists then the other also exists and the two are equal,*

$$g'(t) = f'(a + ty; y). \tag{8.5}$$

In particular, when $t = 0$ we have $g'(0) = f'(a; y)$.

Proof. Forming the difference quotient for g, we have

$$\frac{g(t+h)-g(t)}{h}=\frac{f(a+ty+hy)-f(a+ty)}{h}$$

Letting $h \to 0$ we obtain (8.5).

EXAMPLE 3. Compute $f'(a; y)$ if $f(x) = \|x\|^2$ for all x in $\mathbf{R}^n$.

Solution. We let $g(t) = f(a + ty) = (a + ty) \cdot (a + ty) = a \cdot a + 2ta \cdot y + t^2 y \cdot y$. Therefore $g'(t) = 2a \cdot y + 2ty \cdot y$, so $g'(0) = f'(a; y) = 2a \cdot y$.

A simple corollary of Theorem 8.3 is the mean-value theorem for scalar fields.

THEOREM 8.4. MEAN-VALUE THEOREM FOR DERIVATIVES OF SCALAR FIELDS. *Assume the derivative $f'(a + ty; y)$ exists for each t in the interval $0 \le t \le 1$. Then for some real θ in the open interval $0 < \theta < 1$ we have*

$$f(a+y)-f(a)=f'(z;y), \qquad \textit{where } z = a+\theta y.$$

Proof. Let $g(t) = f(a + ty)$. Applying the one-dimensional mean-value theorem to g on the interval [0, 1] we have

$$g(1) - g(0) = g'(\theta), \qquad \text{where } 0 < \theta < 1.$$

Since $g(1) - g(0) = f(a + y) - f(a)$ and $g'(\theta) = f'(a + \theta y; y)$, this completes the proof.

8.7 Directional derivatives and partial derivatives

In the special case when y is a *unit* vector, that is, when $\|y\| = 1$, the distance between a and $a + hy$ is $|h|$. In this case the difference quotient (8.3) represents the average rate of change of f *per unit distance* along the segment joining a to $a + hy$; the derivative $f'(a; y)$ is called a *directional derivative.*

DEFINITION OF DIRECTIONAL AND PARTIAL DERIVATIVES. *If y is a unit vector, the derivative $f'(a; y)$ is called the directional derivative of f at a in the direction of y. In particular, if $y = e_k$ (the kth unit coordinate vector) the directional derivative $f'(a; e_k)$ is called the partial derivative with respect to e_k and is also denoted by the symbol $D_k f(a)$. Thus,*

$$D_k f(a) = f'(a; e_k).$$

The following notations are also used for the partial derivative $D_k f(a)$:

$$D_k f(a_1, \ldots, a_n), \qquad \frac{\partial f}{\partial x_k}(a_1, \ldots, a_n), \qquad \text{and} \qquad f'_{x_k}(a_1, \ldots, a_n).$$

Sometimes the derivative f'_{x_k} is written without the prime as f_{x_k} or even more simply as f_k.

In $\mathbf{R}^2$ the unit coordinate vectors are denoted by i and j. If a = (a, b) the partial derivatives f '(a; i) and f '(a; j) are also written as

$$\frac{\partial f}{\partial x}(a, b) \quad \text{and} \quad \frac{\partial f}{\partial y}(a, b),$$

respectively. In $\mathbf{R}^3$, if a = (u, b, c) the partial derivatives $D_1f(a)$, $D_2f(a)$, and $D_3f(a)$ are also denoted by

$$\frac{\partial f}{\partial x}(a, b, c), \quad \frac{\partial f}{\partial y}(a, b, c), \quad \text{and} \quad \frac{\partial f}{\partial z}(a, b, c).$$

8.8 Partial, derivatives of higher order

Partial differentiation produces new scalar fields $D_1f, \ldots, D_nf$ from a given scalar field f. The partial derivatives of $D_1f, \ldots, D_nf$ are called ***second-order partial derivatives*** off. For functions of two variables there are four second-order partial derivatives, which are written as follows:

$$D_1(D_1f) = \frac{\partial^2 f}{\partial x^2}, \qquad D_1(D_2f) = \frac{\partial^2 f}{\partial x\, \partial y}, \qquad D_2(D_1f) = \frac{\partial^2 f}{\partial y\, \partial x}, \qquad D_2(D_2f) = \frac{\partial^2 f}{\partial y^2}.$$

Note that $D_1(D_2f)$ means the partial derivative of D_2f with respect to the first variable. We sometimes use the notation $D_{i,j}f$ for the second-order partial derivative $D_i(D_jf)$. For example, $D_{1,2}f = D_1(D_2f)$. In the &notation we indicate the order of derivatives by writing

$$\frac{\partial^2 f}{\partial x\, \partial y} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right).$$

This may or may not be equal to the other mixed partial derivative,

$$\frac{\partial^2 f}{\partial y\, \partial x} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right).$$

In Section 8.23 we shall prove that the two mixed partials $D_1(D_2f)$ and $D_2(D_1f)$ are equal at a point if one of them is continuous in a neighborhood of the point. Section 8.23 also contains an example in which $D_1(D_2f) \neq D_2(D_1f)$ at a point.

8.9 Exercises

1. A scalar field f is defined on $\mathbf{R}^n$ by the equation $f(x)$ = a . x, where a is a constant vector. Compute $f'(x; y)$ for arbitrary x and y.

2. (a) Solve Exercise 1 when $f(x) = \|x\|^4$.
(b) Take $n = 2$ in (a) and find all points (x, y) for which $f'(2i + 3j;\ xi + yj) = 6$.
(c) Take $n = 3$ in (a) and find all points (x, y, z) for which $f'(i + 2j + 3k;\ xi + yj + zk) = 0$.
3. Let $T: \mathbf{R}^n \to \mathbf{R}^n$ be a given linear transformation. Compute the derivative $f'(x; y)$ for the scalar field defined on $\mathbf{R}^n$ by the equation $f(x) = x \cdot T(x)$.

In each of Exercises 4 through 9, compute all first-order partial derivatives of the given scalar field. The fields in Exercises 8 and 9 are defined on $\mathbf{R}^n$.

4. $f(x, y) = x^2 + y^2 \sin (xy)$.
5. $f(x, y) = \sqrt{x^2 + y^2}$.
6. $f(x, y) = \dfrac{x}{\sqrt{x^2 + y^2}}$, $(x, y) \neq (0, 0)$.
7. $f(x, y) = \dfrac{x + y}{x - y}$, $x \neq y$.
8. $f(x) = a \cdot x$, a fixed.
9. $f(x) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j$, $a_{ij} = a_{ji}$.

In each of Exercises 10 through 17, compute all first-order partial derivatives. In each of Exercises 10, 11, and 12 verify that the mixed partials $D_1(D_2 f)$ and $D_2(D_1 f)$ are equal.

10. $f(x, y) = x^4 + y^4 - 4x^2y^2$.
11. $f(x, y) = \log (x^2 + y^2)$, $(x, y) \neq (0, 0)$.
12. $f(x, y) = \dfrac{1}{y} \cos x^2$, $y \neq 0$.
13. $f(x, y) = \tan (x^2/y)$, $y \neq 0$.
14. $f(x, y) = \arctan (y/x)$, $x \neq 0$.
15. $f(x, y) = \arctan \dfrac{x + y}{1 - xy}$, $xy \neq 1$.
16. $f(x, y) = x^{(y^2)}$, $x > 0$.
17. $f(x, y) = \arccos \sqrt{x/y}$, $y \neq 0$.

18. Let $v(r, t) = t^n e^{-r^2/(4t)}$. Find a value of the constant n such that v satisfies the following equation:

$$\frac{\partial v}{\partial t} = \frac{1}{r^2} \frac{\partial}{\partial r}\left(r^2 \frac{\partial v}{\partial r}\right).$$

19. Given $z = u(x, y)e^{ax+by}$ and $\partial^2 u/(\partial x\, \partial y) = 0$. Find values of the constants a and b such that

$$\frac{\partial^2 z}{\partial x \partial y} - \frac{\partial z}{\partial x} - \frac{\partial z}{\partial y} + z = 0.$$

20. (a) Assume that $f'(x; y) = 0$ for every x in some n-ball $B(a)$ and for every vector y. Use the mean-value theorem to prove that f is constant on $B(a)$.
(b) Suppose that $f'(x; y) = 0$ for a *fixed* vector y and for every x in $B(a)$. What can you conclude about f in this case?
21. A set S in $\mathbf{R}^n$ is called convex if for every pair of points a and b in S the line segment from a to b is also in S; in other words, $ta + (1 - t)b \in S$ for each t in the interval $0 \leq t \leq 1$.
(a) Prove that every n-ball is convex.
(b) If $f'(x; y) = 0$ for every x in an open convex set S and for every y in $\mathbf{R}^n$, prove that f is constant on S.
22. (a) Prove that there is no scalar field f such that $f'(a; y) > 0$ for a fixed vector a and every nonzero vector y.
(b) Give an example of a scalar field f such that $f'(x; y) > 0$ for a fixed vector y and every vector x.

8.10 Directional derivatives and continuity

In the one-dimensional theory, existence of the derivative of a function f at a point implies continuity at that point. This is easily proved by choosing an $h \neq 0$ and writing

$$f(a + h) - f(a) = \frac{f(a + h) - f(a)}{h} \cdot h .$$

As $h \to 0$ the right side tends to the limit $f'(a) \cdot 0 = 0$ and hence $f(a + h) \to f(a)$. This shows that the existence of $f'(a)$ implies continuity of f at a.

Suppose we apply the same argument to a general scalar field. Assume the derivative $f'(\mathbf{a}; \mathbf{y})$ exists for some $\mathbf{y}$. Then if $h \neq 0$ we can write

$$f(\mathbf{a} + h\mathbf{y}) - f(\mathbf{a}) = \frac{f(\mathbf{a} + h\mathbf{y}) - f(\mathbf{a})}{h} \cdot h .$$

As $h \to 0$ the right side tends to the limit $f'(\mathbf{a}; \mathbf{y}) \cdot 0 = 0$; hence the existence of $f'(\mathbf{a}; \mathbf{y})$ for a given $\mathbf{y}$ implies that

$$\lim_{h \to 0} f(\mathbf{a} + h\mathbf{y}) = f(\mathbf{a})$$

for the same $\mathbf{y}$. This means that $f(\mathbf{x}) \to f(\mathbf{a})$ as $\mathbf{x} \to \mathbf{a}$ *along a straight line through* $\mathbf{a}$ *having the direction* $\mathbf{y}$. If $f'(\mathbf{a}; \mathbf{y})$ exists for *every* vector $\mathbf{y}$, then $f(\mathbf{x}) \to f(\mathbf{a})$ as $\mathbf{x} \to \mathbf{a}$ along every line through $\mathbf{a}$. This seems to suggest that f is continuous at $\mathbf{a}$. Surprisingly enough, this conclusion need not be true. The next example describes a scalar field which has a directional derivative in every direction at $\mathbf{0}$ but which is not continuous at $\mathbf{0}$.

EXAMPLE. Let f be the scalar field defined on $\mathbf{R}^2$ as follows:

$$f(x, y) = \frac{xy^2}{x^2 + y^4} \quad \text{if } x \neq 0, \quad f(0, y) = 0.$$

Let $\mathbf{a} = (0, 0)$ and let $\mathbf{y} = (a, b)$ be any vector. If $a \neq 0$ and $h \neq 0$ we have

$$\frac{f(\mathbf{a} + h\mathbf{y}) - f(\mathbf{a})}{h} = \frac{f(h\mathbf{y}) - f(\mathbf{O})}{h} = \frac{f(ha, hb)}{h} = \frac{ab^2}{a^2 + h^2b^4} .$$

Letting $h \to 0$ we find $f'(\mathbf{0}; \mathbf{y}) = b^2/a$. If $\mathbf{y} = (0, b)$ we find, in a similar way, that $f'(\mathbf{O}; \mathbf{y}) = 0$. Therefore $f'(\mathbf{O}; \mathbf{y})$ exists for all directions $\mathbf{y}$. Also, $f(\mathbf{x}) \to 0$ as $\mathbf{x} \to \mathbf{0}$ along any straight line through the origin. However, at each point of the parabola $x = y^2$ (except at the origin) the function f has the value $\frac{1}{2}$. Since such points exist arbitrarily close to the origin and since $f(\mathbf{0}) = 0$, the function f is not continuous at $\mathbf{0}$.

The foregoing example shows that the existence of *all* directional derivatives at a point fails to imply continuity at that point. For this reason, directional derivatives are a somewhat unsatisfactory extension of the one-dimensional concept of derivative. A more suitable generalization exists which implies continuity and, at the same time, permits us to extend the principal theorems of one-dimensional derivative theory to the higher dimensional case. This is called the *total derivative.*

8.11 The total derivative

We recall that in the one-dimensional case a function f with a derivative at a can be approximated near a by a linear Taylor polynomial. If $f'(a)$ exists we let $E(a, h)$ denote the difference

$$E(a, h) = \frac{f(a+h) - f(a)}{h} - f'(a) \qquad \text{if } h \neq 0, \tag{8.6}$$

and we define $E(a, 0) = 0$. From (8.6) we obtain the formula

$$f(a + h) = f(a) + f'(a)h + hE(a, h),$$

an equation which holds also for $h = 0$. This is the first-order Taylor formula for approximating $f(a + h) - f(a)$ by $f'(a)h$. The error committed is $hE(a, h)$. From (8.6) we see that $E(a, h) \to 0$ as $h \to 0$. Therefore the error $hE(a, h)$ is of smaller order than h for small h.

This property of approximating a differentiable function by a linear function suggests a way of extending the concept of differentiability to the higher-dimensional case.

Let $f: S \to \mathbf{R}$ be a scalar field defined on a set S in $\mathbf{R}^n$. Let a be an interior point of S, and let $B(\boldsymbol{a}; r)$ be an n-ball lying in S. Let $\boldsymbol{v}$ be a vector with $\|\boldsymbol{v}\| < r$, so that $\boldsymbol{a} + \boldsymbol{v} \in B(\boldsymbol{a}; r)$.

DEFINITION OF A DIFFERENTIABLE SCALAR FIELD. *We say that f is differentiable at $\boldsymbol{a}$ if there exists a linear transformation*

$$T_{\boldsymbol{a}}: \mathbf{R}^n \to \mathbf{R}$$

from $\mathbf{R}^n$ to $\mathbf{R}$, and a scalar function $E(\boldsymbol{a}, \boldsymbol{v})$ such that

$$f(\boldsymbol{a} + \boldsymbol{v}) = f(\boldsymbol{a}) + T_{\boldsymbol{a}}(\boldsymbol{v}) + \|\boldsymbol{v}\|\, E(\boldsymbol{a}, \boldsymbol{v}), \tag{8.7}$$

for $\|\boldsymbol{v}\| < r$, where $E(\boldsymbol{a}, \boldsymbol{v}) \to 0$ as $\|\boldsymbol{v}\| \to 0$. The linear transformation $T_{\boldsymbol{a}}$ is called the total derivative of f at $\boldsymbol{a}$.

Note: The total derivative $T_{\boldsymbol{a}}$ is a linear transformation, not a number. The function value $T_{\boldsymbol{a}}(\boldsymbol{v})$ is a real number; it is defined for every point $\boldsymbol{v}$ in $\mathbf{R}^n$. The total derivative was introduced by W. H. Young in 1908 and by M. Fréchet in 1911 in a more general context.

Equation (8.7), which holds for $\|\boldsymbol{v}\| < r$, is called a *first-order Taylor formula* for $f(\boldsymbol{a} + \boldsymbol{v})$. It gives a linear approximation, $T_{\boldsymbol{a}}(\boldsymbol{v})$, to the difference $f(\boldsymbol{a} + \boldsymbol{v}) - f(\boldsymbol{a})$. The error in the approximation is $\|\boldsymbol{v}\|\, E(\boldsymbol{a}, \boldsymbol{v})$, a term which is of smaller order than $\|\boldsymbol{v}\|$ as $\|\boldsymbol{v}\| \to 0$; that is, $E(\boldsymbol{a}, \boldsymbol{v}) = o(\|\boldsymbol{v}\|)$ as $\|\boldsymbol{v}\| \to 0$.

The next theorem shows that if the total derivative exists it is unique. It also tells us how to compute $T_{\boldsymbol{a}}(\boldsymbol{y})$ for every $\boldsymbol{y}$ in $\mathbf{R}^n$.

THEOREM 8.5. *Assume f is differentiable at a with total derivative T_a. Then the derivative $f'(a; y)$ exists for every y in $\mathbf{R}^n$ and we have*

$$T_a(y) = f'(a; y). \tag{8.8}$$

Moreover, $f'(a; y)$ is a linear combination of the components of y. In fact, if $y = (y_1, \ldots, y_n)$, we have

$$f'(a; y) = \sum_{k=1}^{n} D_k f(a) y_k. \tag{8.9}$$

Proof. Equation (8.8) holds trivially if $y = 0$ since both $T_a(O) = 0$ and $f'(a; O) = 0$. Therefore we can assume that $y \neq 0$.

Since f is differentiable at a we have a Taylor formula,

$$f(a + v) = f(a) + T_a(v) + \|v\| E(a, v) \tag{8.10}$$

for $\|v\| < r$ for some $r > 0$, where $E(a, v) \to 0$ as $\|v\| \to 0$. In this formula we take $v = hy$, where $h \neq 0$ and $|h| \|y\| < r$. Then $\|v\| < r$. Since T_a is linear we have $T_a(v) = T_a(hy) = hT_a(y)$. Therefore (8.10) gives us

$$\frac{f(a + hy) - f(a)}{h} = T_a(y) + \frac{|h| \|y\|}{h} E(a, v). \tag{8.11}$$

Since $\|v\| \to 0$ as $h \to 0$ and since $|h|/h = \pm 1$, the right-hand member of (8.11) tends to the limit $T_a(y)$ as $h \to 0$. Therefore the left-hand member tends to the same limit. This proves (8.8).

Now we use the linearity of T_a to deduce (8.9). If $y = (y_1, \ldots, y_n)$ we have $y = \sum_{k=1}^{n} y_k e_k$, hence

$$T_a(y) = T_a\left(\sum_{k=1}^{n} y_k e_k\right) = \sum_{k=1}^{n} y_k T_a(e_k) = \sum_{k=1}^{n} y_k f'(a; e_k) = \sum_{k=1}^{n} y_k D_k f(a).$$

8.12 The gradient of a scalar field

The formula in Theorem 8.5, which expresses $f'(a; y)$ as a linear combination of the components of y, can be written as a dot product,

$$f'(a; y) = \sum_{k=1}^{n} D_k f(a) y_k = \nabla f(a) \cdot y,$$

where $\nabla f(a)$ is the vector whose components are the partial derivatives of f at a,

$$\nabla f(a) = (D_1 f(a), \ldots, D_n f(a)).$$

This is called the *gradient* of f. The gradient ∇f is a vector field defined at each point a where the partial derivatives $D_1 f(a), \ldots, D_n f(a)$ exist. We also write grad f for ∇f. The symbol ∇ is pronounced "del."

The first-order Taylor formula (8.10) can now be written in the form

$$f(a + v) = f(a) + \nabla f(a) \cdot v + \|v\| E(a, v), \tag{8.12}$$

where $E(a, v) \to 0$ as $\|v\| \to 0$. In this form it resembles the one-dimensional Taylor formula, with the gradient vector $\nabla f(a)$ playing the role of the derivative $f'(a)$.

From the Taylor formula we can easily prove that differentiability implies continuity.

THEOREM 8.6. *If a scalar field f is differentiable at a, then f is continuous at a.*

Proof. From (8.12) we have

$$|f(a + v) - f(a)| = |\nabla f(a) \cdot v + \|v\| E(a, v)|.$$

Applying the triangle inequality and the Cauchy-Schwarz inequality we find

$$0 \le |f(a + v) - f(a)| \le \|\nabla f(a)\| \|v\| + \|v\| |E(a, v)|.$$

This shows that $f(a + v) \to f(a)$ as $\|v\| \to 0$, so f is continuous at a.

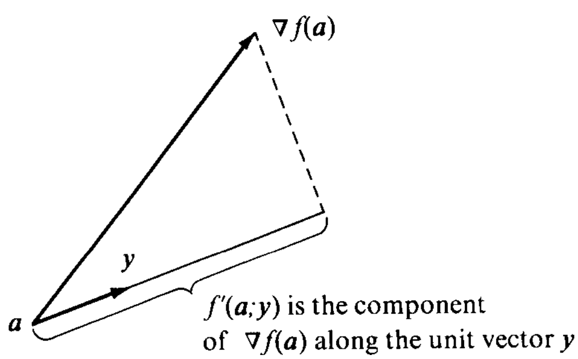

FIGURE 8.5 Geometric relation of the directional derivative to the gradient vector.

When y is a *unit* vector the directional derivative $f'(a; y)$ has a simple geometric relation to the gradient vector. Assume that $\nabla f(a) \neq 0$ and let θ denote the angle between y and $\nabla f(a)$. Then we have

$$f'(a; y) = \nabla f(a) \cdot y = \|\nabla f(a)\| \|y\| \cos \theta = \|\nabla f(a)\| \cos \theta.$$

This shows that the directional derivative is simply the component of the gradient vector in the direction of y. Figure 8.5 shows the vectors $\nabla f(a)$ and y attached to the point a. The derivative is largest when $\cos \theta = 1$, that is, when y has the same direction as $\nabla f(a)$. In other words, at a given point a, the scalar field undergoes its maximum rate of change in the direction of the gradient vector; moreover, this maximum is equal to the length of the gradient vector. When $\nabla f(a)$ is orthogonal to y, the directional derivative $f'(a; y)$ is 0.

In 2-space the gradient vector is often written as

$$\nabla f(x, y) = \frac{\partial f(x, y)}{\partial x} \boldsymbol{i} + \frac{\partial f(x, y)}{\partial y} \boldsymbol{j}.$$

In 3-space the corresponding formula is

$$\nabla f(x, y, z) = \frac{\partial f(x, y, z)}{\partial x} \boldsymbol{i} + \frac{\partial f(x, y, z)}{\partial y} \boldsymbol{j} + \frac{\partial f(x, y, z)}{\partial z} \boldsymbol{k}$$

8.13 A sufficient condition for differentiability

Iffis differentiable at $\boldsymbol{a}$, then all partial derivatives $D_1 f(\boldsymbol{a}), \ldots, D_n f(\boldsymbol{a})$ exist. However, the existence of all these partials does not necessarily imply that f is differentiable at $\boldsymbol{a}$. A counter example is provided by the function

$$f(x, y) = \frac{xy^2}{x^2 + y^4} \quad \text{if } x \neq 0, \quad f(0, y) = 0,$$

discussed in Section 8.10. For this function, both partial derivatives $D_1 f(\boldsymbol{O})$ and $D_2 f(\boldsymbol{O})$ exist but f is not continuous at 0, hence f cannot be differentiable at 0.

The next theorem shows that the existence of *continuous* partial derivatives at a point implies differentiability at that point.

THEOREM 8.7. A SUFFICIENT CONDITION FOR DIFFERENTIABILITY. *Assume that thepartial derivatives $D_1 f, \ldots, D_n f$ exist in some n-ball B(a) and are continuous at a. Then f is differentiable at a.*

> *Note:* A scalar field satisfying the hypothesis of Theorem 8.7 is said to be *continuously differentiable* at $\boldsymbol{a}$.

Proof. The only candidate for $T_{\boldsymbol{a}}(\boldsymbol{v})$ is $\nabla f(\boldsymbol{a}) \cdot \boldsymbol{v}$. We will show that

$$f(\boldsymbol{a} + \boldsymbol{v}) - f(\boldsymbol{a}) = \nabla f(\boldsymbol{a}) \cdot \boldsymbol{v} + \|\boldsymbol{v}\| E(\boldsymbol{a}, \boldsymbol{v}),$$

where $E(\boldsymbol{a}, \boldsymbol{v}) \to 0$ as $\|\boldsymbol{v}\| \to 0$. This will prove the theorem.

Let $\lambda = \|\boldsymbol{v}\|$. Then $\boldsymbol{v} = \lambda \boldsymbol{u}$, where $\|\boldsymbol{u}\| = 1$. We keep λ small enough so that $\boldsymbol{a} + \boldsymbol{v}$ lies in the ball $B(\boldsymbol{a})$ in which the partial derivatives $D_1 f, \ldots, D_n f$ exist. Expressing $\boldsymbol{u}$ in terms of its components we have

$$\boldsymbol{u} = u_1 \boldsymbol{e}_1 + \cdots + u_n \boldsymbol{e}_n,$$

where $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ are the unit coordinate vectors. Now we write the difference $f(\boldsymbol{a} + \boldsymbol{v}) - f(\boldsymbol{a})$ as a telescoping sum,

$$(8.13) \qquad f(\boldsymbol{a} + \boldsymbol{v}) - f(\boldsymbol{a}) = f(\boldsymbol{a} + \lambda \boldsymbol{u}) - f(\boldsymbol{a}) = \sum_{k=1}^{n} \{f(\boldsymbol{a} + \lambda \boldsymbol{v}_k) - f(\boldsymbol{a} + \lambda \boldsymbol{v}_{k-1})\},$$

where $v_0, v_1, \ldots, v_n$ are any vectors in $\mathbf{R}^n$ such that $v_0 = 0$ and $v_n = u$. We choose these vectors so they satisfy the recurrence relation $v_k = v_{k-1} + u_k e_k$. That is, we take

$$v_0 = 0\,, \quad v_1 = u_1 e_1\,, \quad v_2 = u_1 e_1 + u_2 e_2\,, \quad \ldots, \quad v_n = u_1 e_1 + \cdots + u_n e_n\,.$$

Then the kth term of the sum in (8.13) becomes

$$f(a + \lambda v_{k-1} + \lambda u_k e_k) - f(a + \lambda v_{k-1}) = f(b_k + \lambda u_k e_k) - f(b_k)\,,$$

where b_k = a + λv_{k-1} . The two points b_k and $b_k + \lambda u_k e_k$ differ only in their kth component. Therefore we can apply the mean-value theorem of differential calculus to write

$$f(b_k + \lambda u_k e_k) - f(b_k) = \lambda u_k\, D_k f(c_k)\,, \tag{8.14}$$

where c_k lies on the line segment joining b_k to $b_k + \lambda u_k e_k$. Note that $b_k \to$ a and hence $c_k \to a$ as $\lambda \to 0$.

Using (8.14) in (8.13) we obtain

$$f(a + v) - f(a) = \lambda \sum_{k=1}^{n} D_k f(c_k) u_k\,.$$

But $\nabla f(a) \cdot v = \lambda$ V'(u) . $u = \lambda \sum_{k=1}^{n} D_k f(a) u_k$, *so*

$$f(a + v) - f(a) - \nabla f(a) \cdot v = \lambda \sum_{h=1}^{n} \{D_k f(c_k) - D_k f(a)\} u_k = \|v\|\, E(a, v)\,,$$

where

$$E(a, v) = \sum_{k=1}^{n} \{D_k f(c_k) - D_k f(a)\} u_k\,.$$

Since $c_k \to$ a as $\|v\| \to 0$, and since each partial derivative $D_k f$ is continuous at a, we see that $E(a, v) \to 0$ as $\|v\| \to 0$. This completes the proof.

8.14 Exercises

1. Find the gradient vector at each point at which it exists for the scalar fields defined by the following equations :
   (a) $f(x, y) = x^2 + y^2 \sin(xy)$.
   (b) $f(x, y) = e^x \cos y$.
   (c) $f(x, y, z) = x^2 y^3 z^4$.
   (d) $f(x, y, z) = x^2 - y^2 + 2z^2$.
   (e) $f(x, y, z) = \log(x^2 + 2y^2 - 3z^2)$.
   (f) $f(x, y, z) = x^{y^z}$.
2. Evaluate the directional derivatives of the following scalar fields for the points and directions given :
   (a) $f(x, y, z) = x^2 + 2y^2 + 3z^2$ at (1, 1, 0) in the direction of i -j + $2k$.
   (b) $f(x, y, z)$ = (x/y)" at (1, 1, 1) in the direction of $2i + j - k$.
3. Find the points (x, y) and the directions for which the directional derivative of $f(x, y) = 3x^2 + y^2$ has its largest value, if (x, y) is restricted to be on the circle $x^2 + y^2 = 1$.
4. A differentiable scalar fieldfhas, at the point (1, 2), directional derivatives +2 in the direction toward (2, 2) and -2 in the direction toward (1, 1). Determine the gradient vector at (1, 2) and compute the directional derivative in the direction toward (4, 6).
5. Find values of the constants a, b, and c such that the directional derivative of $f(x, y, z) = axy^2 + byz + cz^2x^3$ at the point (1, 2, $-$ 1) has a maximum value of 64 in a direction parallel to the z-axis.

6. Given a scalar field differentiable at a point a in $\mathbf{R}^2$. Suppose that $f'(a; y) = 1$ and $f'(a; z) = 2$, where $y = 2i + 3j$ and $z = i + j$. Make a sketch showing the set of all points (x, y) for which $f'(a; xi + yj) = 6$. Also, calculate the gradient V'(a).
7. Let f andg denote scalar fields that are differentiable on an open set S. Derive the following properties of the gradient:
   (a) grad $f = 0$ iff is constant on S.
   (b) grad $(f + g)$ = gradf + gradg .
   (c) grad (cf) = c gradfif c is a constant.
   (d) grad $(fg) = f \operatorname{grad} g + g \operatorname{grad} f$.
   (e) grad $\frac{f}{g} = \frac{g \operatorname{grad} f - f \operatorname{grad} g}{g^2}$ at points at which $g \neq 0$.
8. In $\mathbf{R}^3$ let $r(x, y, z) = xi + yj + zk$, and let $r(x, y, z) = \|r(x, y, z)\|$.
   (a) Show that $\nabla r(x, y, z)$ is a unit vector in the direction of $r(x, y, z)$.
   (b) Show that $\nabla(r^n) = nr^{n-2}r$ if n is a positive integer. *[Hint:* Use Exercise 7(d).]
   (c) Is the formula of part (b) valid when n is a negative integer or zero?
   (d) Find a scalar field f such that $\nabla f = r$.
9. Assume f is differentiable at each point of an n-ball $B(a)$. If $f'(x; y) = 0$ for n independent vectors $y_1, \ldots, y_n$ and for every x in B(a), prove that f is constant on B(a).
10. Assume f is differentiable at each point of an n-ball B(a).
    (a) If $\nabla f(x) = 0$ for every x in B(u), prove that f is constant on B(a).
    (b) If $f(x) \leq f(a)$ for all x in B(u), prove that $\nabla f(a) = 0$.
11. Consider the following six statements about a scalar field $f: S \to \mathbf{R}$, where $S \subseteq \mathbf{R}^n$ and a is an interior point of S.
    (a) f is continuous at a.
    (b) f is differentiable at a.
    (c) $f'(a; y)$ exists for every y in $\mathbf{R}^n$.
    (d) All the first-order partial derivatives off exist in a neighborhood of a and are continuous at a.
    (e) $\nabla f(a) = 0$.
    (f) $f(x) = \|x - a\|$ for all x in $\mathbf{R}^n$.

In a table like the one shown here, mark T in the appropriate square if the statement in row (x) always implies the statement in column (y). For example, if (a) always implies (b), mark T in the second square of the first row. The main diagonal has already been filled in for you.

	a	b	c	d	e	f
a	T					
b		T				
c			T			
d				T		
e					T	
f						T

8.15 A chain rule for derivatives of scalar fields

In one-dimensional derivative theory, the chain rule enables us to compute the derivative of a composite function $g(t) = f[r(t)]$ by the formula

$$g'(t) = f'[r(t)] \cdot r'(t).$$