Proof. Forming the difference quotient for g, we have

$$\frac{g(t+h)-g(t)}{h} = \frac{f(\boldsymbol{a}+t\boldsymbol{y}+h\boldsymbol{y})-f(\boldsymbol{a}+t\boldsymbol{y})}{h}$$

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

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

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

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 $f'(\boldsymbol{a}+t\boldsymbol{y}; \boldsymbol{y})$ 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*

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

Proof. Let $g(t) = f(\boldsymbol{a}+t\boldsymbol{y})$. 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 $g(1) - g(0) = f(\boldsymbol{a}+\boldsymbol{y}) - f(\boldsymbol{a})$ and $g'(\theta) = f'(\boldsymbol{a}+\theta\boldsymbol{y}; \boldsymbol{y})$, this completes the proof.

8.7 Directional derivatives and partial derivatives

In the special case when $\boldsymbol{y}$ is a *unit* vector, that is, when $\|\boldsymbol{y}\| = 1$, the distance between $\boldsymbol{a}$ and $\boldsymbol{a} + h\boldsymbol{y}$ 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 $\boldsymbol{a}$ to $\boldsymbol{a} + h\boldsymbol{y}$; the derivative $f'(\boldsymbol{a}; \boldsymbol{y})$ is called a *directional derivative.*

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

$$D_k f(\boldsymbol{a}) = f'(\boldsymbol{a}; \boldsymbol{e}_k).$$

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

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

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

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

respectively. In $\mathbf{R}^3$, if a $=$ (u, b, c) the partial derivatives $D_1 f(\boldsymbol{a})$, $D_2 f(\boldsymbol{a})$, and $D_3 f(\boldsymbol{a})$ are also denoted by

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

8.8 Partial, derivatives of higher order

Partial differentiation produces new scalar fields $D_1 f, \ldots, D_n f$ from a given scalar field f. The partial derivatives of $D_1 f, \ldots, D_n f$ are called ***second-order partial derivatives*** off. For functions of two variables there are four second-order partial derivatives, which are written as follows:

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

Note that $D_1(D_2 f)$ means the partial derivative of $D_2 f$ with respect to the first variable. We sometimes use the notation $D_{i,j} f$ for the second-order partial derivative $D_i(D_j f)$. For example, $D_{1,2} f = D_1(D_2 f)$. 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 $D_1(D_2 f)$ and $D_2(D_1 f)$ 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 $D_1(D_2 f) \neq D_2(D_1 f)$ at a point.

8.9 Exercises

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

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

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. $f(\boldsymbol{x}) = \boldsymbol{a} \cdot \boldsymbol{x}$, $\boldsymbol{a}$ fixed.
9. $f(\boldsymbol{x}) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}x_ix_j$, $a_{ij} = a_{ji}$.

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 $D_1(D_2f)$ and $D_2(D_1f)$ 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 $f'(\boldsymbol{x}; \boldsymbol{y}) = 0$ for every $\boldsymbol{x}$ in some n-ball $B(\boldsymbol{a})$ and for every vector $\boldsymbol{y}$. Use the mean-value theorem to prove that f is constant on $B(\boldsymbol{a})$.
(b) Suppose that $f'(\boldsymbol{x}; \boldsymbol{y}) = 0$ for a *fixed* vector $\boldsymbol{y}$ and for every $\boldsymbol{x}$ in $B(\boldsymbol{a})$. 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 $\boldsymbol{a}$ and $\boldsymbol{b}$ in S the line segment from $\boldsymbol{a}$ to $\boldsymbol{b}$ is also in S; in other words, $t\boldsymbol{a} + (1 - t)\boldsymbol{b} \in S$ for each t in the interval $0 \leq t \leq 1$.
(a) Prove that every n-ball is convex.
(b) If $f'(\boldsymbol{x}; \boldsymbol{y}) = 0$ for every $\boldsymbol{x}$ in an open convex set S and for every $\boldsymbol{y}$ in $\mathbf{R}^n$, prove that f is constant on S.
22. (a) Prove that there is no scalar field f such that $f'(\boldsymbol{a}; \boldsymbol{y}) > 0$ for a fixed vector $\boldsymbol{a}$ and every nonzero vector $\boldsymbol{y}$.
(b) Give an example of a scalar field f such that $f'(\boldsymbol{x}; \boldsymbol{y}) > 0$ for a fixed vector $\boldsymbol{y}$ and every vector $\boldsymbol{x}$.

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 $f'(\boldsymbol{a}; \boldsymbol{y})$ exists for some $\boldsymbol{y}$. Then if $h \neq 0$ we can write

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

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

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

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

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

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

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

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

Letting $h \to 0$ we find $f'(\boldsymbol{O}; \boldsymbol{y}) = b^2/a$. If $\boldsymbol{y} = (0, b)$ we find, in a similar way, that $f'(\boldsymbol{O}; \boldsymbol{y}) = 0$. Therefore $f'(\boldsymbol{O}; \boldsymbol{y})$ exists for all directions $\boldsymbol{y}$. Also, $f(\boldsymbol{x}) \to 0$ as $\boldsymbol{x} \to \boldsymbol{O}$ 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 $f(\boldsymbol{O}) = 0$, the function f is not continuous at $\boldsymbol{O}$.

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. Iff '(a) exists we let $E(a, h)$ denote the difference

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

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 R”. Let a be an interior point of S, and let $B(a; r)$ be an n-ball lying in S. Let v be a vector with $\|v\| < r$, so that $a + v \in B(a; r)$.

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

$$T_a: \mathbf{R}^n \to \mathbf{R}$$

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

$$(8.7) \qquad f(a + v) = f(a) + T_a(v) + \|v\| \, E(a, v),$$

for $\|v\| < r$, where $E(a, v) \to 0$ as $\|v\| \to 0$. The linear transformation T_a is called the total derivative off at a.

Note: The total derivative T_a is a linear transformation, not a number. The function value $T_a(v)$ is a real number; it is defined for every point 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 $\|v\| < r$, is called a *first-order Taylorformula* forf **(a + u).** It gives a linear approximation, *T,(u),* to the difference $f(a + v)$ -f(u). The error in the approximation is $\|v\| \, E(a, v)$, a term which is of smaller order than $\|v\|$ as $\|v\| \to 0$; that is, $E(a, v) = o(\|v\|)$ as $\|v\| \to 0$.

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

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

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

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

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

Proof. Equation (8.8) holds trivially if y = 0 since both *T,(O)* = 0 andf'(a; *O*) = 0. Therefore we can assume that y $\neq$ 0.

Since f is differentiable at $\boldsymbol{a}$ we have a Taylor formula,

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

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

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

Since $\|\boldsymbol{v}\| \to 0$ as $h \to 0$ and since $|h|/h = \pm 1$, the right-hand member of (8.11) tends to the limit *T,(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 = $\sum_{k=1}^{n} y_k \boldsymbol{e}_k$, hence

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

8.12 The gradient of a scalar field

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

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

where ∇f *(a)* is the vector whose components are the partial derivatives off at ***a,***

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

This is called the ***gradient*** off The gradient ∇f is a vector field defined at each point $\boldsymbol{a}$ where the partial derivatives $D_1 f(\boldsymbol{a}), \ldots, D_n f(\boldsymbol{a})$ exist. We also write grad f for Vf. The symbol V 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.

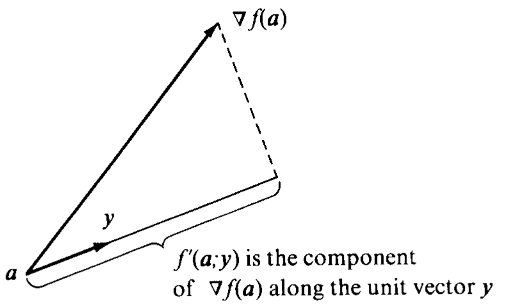

FIGURE 8.5 Geometric relation of the directional derivative to the gradient vector.

When y is a ***unit*** vector the directional derivative $f'(a; y)$ has a simple geometric relation to the gradient vector. Assume that $\nabla f(a) \ne 0$ 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

If f is 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

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

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 the partial derivatives $D_1 f, \ldots, D_n f$ exist in some n-ball $B(\boldsymbol{a})$ and are continuous at $\boldsymbol{a}$. Then f is differentiable at $\boldsymbol{a}$.*

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,

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

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 $b_k = a + \lambda v_{k-1}$. 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) $f(x, y, z) = (x/y)^z$ 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 ∇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) ∇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

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

This section provides an extension of the formula when f is replaced by a scalar field defined on a set in n-space and r is replaced by a vector-valued function of a real variable with values in the domain off.

In a later section we further extend the formula to cover the case in which both f and r are vector fields.

It is easy to conceive of examples in which the composition of a scalar field and a vector field might arise. For instance, suppose $f(\boldsymbol{x})$ measures the temperature at a point $\boldsymbol{x}$ of a solid in 3-space, and suppose we wish to know how the temperature changes as the point $\boldsymbol{x}$ varies along a curve C lying in the solid. If the curve is described by a vector-valued function $\boldsymbol{r}$ defined on an interval $[a, b]$, we can introduce a new function g by means of the formula

$$g(t) = f[\boldsymbol{r}(t)] \qquad \text{if} \quad a \leq t \leq b.$$

This composite function g expresses the temperature as a function of the parameter t, and its derivative $g'(t)$ measures the rate of change of the temperature along the curve. The following extension of the chain rule enables us to compute the derivative $g'(t)$ without determining $g(t)$ explicitly.

THEOREM 8.8. CHAIN RULE. *Let f be a scalar field defined on an open set S in $\mathbf{R}^n$, and let r be a vector-valuedfunction which maps an interval J from $\mathbf{R}^1$ into S. Define the composite function $g = f \circ r$ on J by the equation*

$$g(t) = f[\boldsymbol{r}(t)] \qquad \text{if} \quad t \in J.$$

Let t be a point in J at which $\boldsymbol{r}'(t)$ exists and assume that f is differentiable at r(t). Then g'(t) exists and is equal to the dot product

$$g'(t) = \nabla f(\boldsymbol{a}) \cdot \boldsymbol{r}'(t), \qquad \text{where } \boldsymbol{a} = \boldsymbol{r}(t). \tag{8.15}$$

Proof. Let $\boldsymbol{a} = \boldsymbol{r}(t)$, where t is a point in J at which $\boldsymbol{r}'(t)$ exists. Since S is open there is an n-ball $B(\boldsymbol{a})$ lying in S. We take $h \neq 0$ but small enough so that $\boldsymbol{r}(t + h)$ lies in $B(\boldsymbol{a})$, and we let $\boldsymbol{y} = \boldsymbol{r}(t + h) - \boldsymbol{r}(t)$. Note that $\boldsymbol{y} \to 0$ as $h \to 0$. Now we have

$$g(t + h) - g(t) = f[\boldsymbol{r}(t + h)] - f[\boldsymbol{r}(t)] = f(\boldsymbol{a} + \boldsymbol{y}) - f(\boldsymbol{a})$$

Applying the first-order Taylor formula for f we have

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

where $E(\boldsymbol{a}, \boldsymbol{y}) \to 0$ as $\|\boldsymbol{y}\| \to 0$. Since $\boldsymbol{y} = \boldsymbol{r}(t + h) - \boldsymbol{r}(t)$ this gives us

$$\frac{g(t + h) - g(t)}{h} = \nabla f(\boldsymbol{a}) \cdot \frac{\boldsymbol{r}(t + h) - \boldsymbol{r}(t)}{h} + \frac{\|\boldsymbol{r}(t + h) - \boldsymbol{r}(t)\|}{h} E(\boldsymbol{a}, \boldsymbol{y}).$$

Letting $h \to 0$ we obtain (8.15).