EXAMPLE 1. *Directional derivative along a curve.* When the function r describes a curve C, the derivative r' is the velocity vector (tangent to the curve) and the derivative g' in Equation (8.15) is the derivative off with respect to the velocity vector, assuming that $r' \neq 0$. If $T(t)$ is a unit vector in the direction of $r'(t)$ (Tis the unit tangent vector), the dot product Vf $[r(t)] \cdot T(t)$ is called the *directional derivative off along the curve C* or *in the direction of* C. For a plane curve we can write

$$T(t) = \cos \alpha(t)\, i + \cos \beta(t)\, j,$$

where $\alpha(t)$ and $\beta(t)$ are the angles made by the vector $T(t)$ and the positive x- and y-axes; the directional derivative off along C becomes

$$\nabla f\ [r(t)] \cdot T(t) = D_1 f[r(t)] \cos a(t) + D_2 f[r(t)] \cos \beta(t).$$

This formula is often written more briefly as

$$\nabla f \cdot T = \frac{\partial f}{\partial x} \cos \alpha + \frac{\partial f}{\partial y} \cos \beta .$$

Some authors write df/ds for the directional derivative $\nabla f \cdot T$. Since the directional derivative along C is defined in terms of T, its value depends on the parametric representation chosen for C. A change of the representation could reverse the direction of T; this, in turn, would reverse the sign of the directional derivative.

EXAMPLE 2. Find the directional derivative of the scalar field $f(x, y) = x^2 - 3xy$ along the parabola $y = x^2 - x + 2$ at the point (1,2).

Solution. At an arbitrary point (x, y) the gradient vector is

$$\nabla f(x, y) = \frac{\partial f}{\partial x} i + \frac{\partial f}{\partial y} i = (2x - 3y)i - 3xj.$$

At the point $(1, 2)$ we have $\nabla f(1, 2) = -4i - 3j$. The parabola can be represented parametrically by the vector equation $r(t) = ti + (t^2 - t + 2)j$. Therefore, $r(1) = i + 2j$, $r'(t) = i + (2t - 1)j$, and $r'(l) = i + j$. For this representation of C the unit tangent vector T(1) is $(i + j)/\sqrt{2}$ and the required directional derivative is $\nabla f(1, 2) \cdot T(1) = -7/\sqrt{2}$.

EXAMPLE 3. Let f be a nonconstant scalar field, differentiable everywhere in the plane, and let c be a constant. Assume the Cartesian equation $f(x, y) = c$ describes a curve C having a tangent at each of its points. Prove that f has the following properties at each point of C:

(a) The gradient vector ∇f is normal to C.
(b) The directional derivative off is zero along C.
(c) The directional derivative off has its largest value in a direction normal to C.

Solution. If $\boldsymbol{T}$ is a unit tangent vector to C, the directional derivative off along C is the dot product $\nabla f \cdot \boldsymbol{T}$. This product is zero if ∇f is perpendicular to $\boldsymbol{T}$, and it has its largest value if ∇f is parallel to $\boldsymbol{T}$. Therefore both statements (b) and (c) are consequences of (a). To prove (a), consider any plane curve Γ with a vector equation of the form $\boldsymbol{r}(t) = X(t)\boldsymbol{i} + Y(t)\boldsymbol{j}$ and introduce the function $g(t) = f[\boldsymbol{r}(t)]$. By the chain rule we have $g'(t) = \nabla f[\boldsymbol{r}(t)] \cdot \boldsymbol{r}'(t)$. When $\Gamma = C$, the function g has the constant value c so $g'(t) = 0$ if $\boldsymbol{r}(t) \in C$. Since $g' = \nabla f \cdot \boldsymbol{r}'$, this shows that ∇f is perpendicular to $\boldsymbol{r}'$ on C; hence ∇f is normal to C.

8.16 Applications to geometry. Level sets. Tangent planes

The chain rule can be used to deduce geometric properties of the gradient vector. Let f be a scalar field defined on a set S in $\mathbf{R}^n$ and consider those points $\boldsymbol{x}$ in S for which $f(\boldsymbol{x})$ has a constant value, say $f(\boldsymbol{x}) = c$. Denote this set by L(c), so that

$$L(c) = \{\boldsymbol{x} \mid \boldsymbol{x} \in S \text{ and } f(\boldsymbol{x}) = c\}.$$

The set $L(c)$ is called a *level set* off. In $\mathbf{R}^2$, $L(c)$ is called a *level curve;* in $\mathbf{R}^3$ it is called a *level surface.*

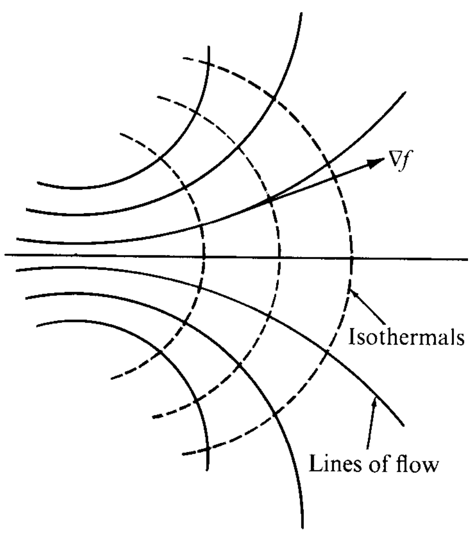

FIGURE 8.6 The dotted curves are isothermals: $f(x, y) = c$. The gradient vector **Vf** points in the direction of the lines of flow.

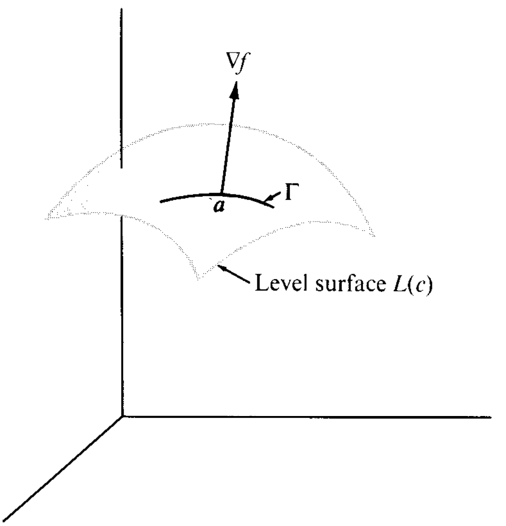

FIGURE 8.7 The gradient vector ∇f is normal to each curve Γ on the level surface $f(x, y, z) = c$.

Families of level sets occur in many physical applications. For example, if $f(x, y)$ represents temperature at (x, y), the level curves off (curves of constant temperature) are called *isothermals.* The flow of heat takes place in the direction of most rapid change in temperature. As was shown in Example 3 of the foregoing section, this direction is normal to the isothermals. Hence, in a thin flat sheet the flow of heat is along a family of curves orthogonal to the isothermals. These are called the *lines of flow*; they are the orthogonal trajectories of the isothermals. Examples are shown in Figure 8.6.

Now consider a scalar fieldfdifferentiable on an open set S in $\mathbf{R}^3$, and examine one of its level surfaces, *L(c)*. Let $\boldsymbol{a}$ be a point on this surface, and consider a curve Γ which lies on the surface and passes through $\boldsymbol{a}$, as suggested by Figure 8.7. We shall prove that the gradient vector V'(a) is normal to this curve at $\boldsymbol{a}$. That is, we shall prove that $\nabla f(\boldsymbol{a})$ is perpendicular to the tangent vector of I' at a. For this purpose we assume that Γ is

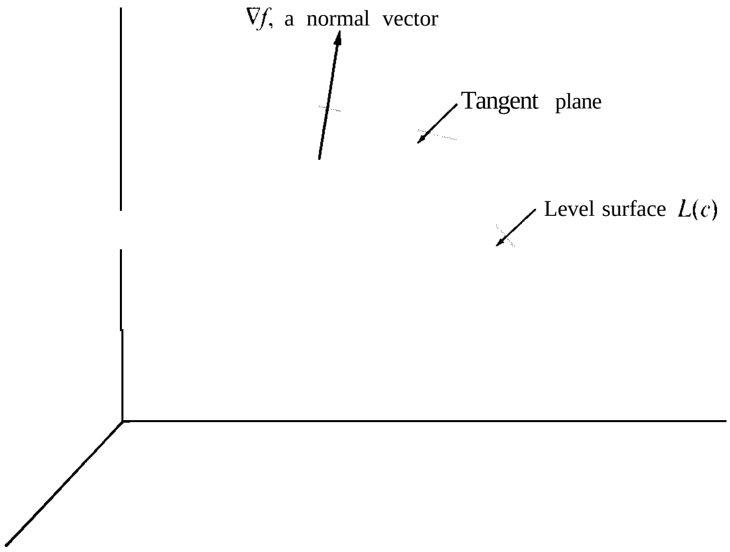

FIGURE 8.8 The gradient vector ∇f is normal to the tangent plane of a level surface $f(x, y, z) = c$.

described parametrically by a differentiable vector-valued function *r* defined on some interval J in $\mathbf{R}^1$. Since Γ lies on the level surface *L(c),* the function *r* satisfies the equation

$$f[\boldsymbol{r}(t)] = c \qquad \text{for all } t \text{ in } J.$$

If g(r) $= f[\boldsymbol{r}(t)]$ for t in J, the chain rule states that

$$g'(t) = \nabla f[\boldsymbol{r}(t)] \cdot \boldsymbol{r}'(t).$$

Since g is constant on J, we have $g'(t) = 0$ on J. In particular, choosing t_1 so that $g(t_1) = \boldsymbol{a}$ we find that

$$\nabla f(\boldsymbol{a}) \cdot \boldsymbol{r}'(t_1) = 0 .$$

In other words, the gradient off at a is perpendicular to the tangent vector $\boldsymbol{r}'(t_1)$, as asserted.

Now. we take a family of curves on the level surface L(c), all passing through the point a. According to the foregoing discussion, the tangent vectors of all these curves are perpendicular to the gradient vector V'(a). If $\nabla f(a)$ is not the zero vector, these tangent vectors determine a plane, and the gradient V'(a) is normal to this plane. (See Figure 8.8.). This particular plane is called the *tangent plane* of the surface $L(c)$ at a.

We know from Volume I that a plane through a with normal vector N consists of all points x in $\mathbf{R}^3$ satisfying $N \cdot (x - a) = 0$. Therefore the tangent plane to the level surface $L(c)$ at a consists of all x in $\mathbf{R}^3$ satisfying

$$\nabla f(a) \cdot (x - a) = 0.$$

To obtain a Cartesian equation for this plane we express x, a, and V'(u) in terms of their components. Writing x = (x, y, z), a $= (x_1, y_1, z_1)$, and

$$\nabla f(a) = D_1 f(a) i + D_2 f(a) j + D_3 f(a) k,$$

we obtain the Cartesian equation

$$D_1 f(a)(x - x_1) + D_2 f(a)(y - y_1) + D_3 f(a)(z - z_1) = 0.$$

A similar discussion applies to scalar fields defined in $\mathbf{R}^2$. In Example 3 of the foregoing section we proved that the gradient vector $\nabla f(a)$ at a point a of a level curve is perpendicular to the tangent vector of the curve at a. Therefore the tangent line of the level curve $L(c)$ at a point a = (x_1, y_1) has the Cartesian equation

$$D_1 f(a)(x - x_1) + D_2 f(a)(y - y_1) = 0.$$

8.17 Exercises

1. In this exercise you may assume the existence and continuity of all derivatives under consideration. The equations $u = f(x, y)$, x = X(t), y = $Y(t)$ define u as a function of t, say $u = F(t)$.

   (a) Use the chain rule to show that

   $$F'(t) = \frac{\partial f}{\partial x} X'(t) + \frac{\partial f}{\partial y} Y'(t),$$

   where $\partial f/\partial x$ and $\partial f/\partial y$ are to be evaluated at [X(t), Y(r)].

   (b) In a similar way, express the second derivative F"(t) in terms of derivatives off, X, and Y. Remember that the partial derivatives in the formula of part (a) are composite functions given by

   $$\frac{\partial f}{\partial x} = D_1 f[X(t), Y(t)], \qquad \frac{\partial f}{\partial y} = D_2 f[X(t), Y(t)].$$

2. Refer to Exercise 1 and compute F'(t) and F"(t) in terms of t for each of the following special cases:

   (a) $f(x, y) = x^2 + y^2$, $X(t) = t$, $Y(t) = t^2$.
   (b) $f(x, y) = e^{xy} \cos (xy^2)$, X(t) = cos t, Y(t) = sin t.
   (c) $f(x, y) = \log [(1 + e^{x^2})/(1 + e^{y^2})]$, X(t) = e, $Y(t)^t = e^{-t}$.

3. In each case, evaluate the directional derivative off for the points and directions specified:
(a) $f(x, y, z) = 3x - 5y + 2z$ at $(2, 2, 1)$ in the direction of the outward normal to the sphere $x^2 + y^2 + z^2 = 9$.
(b) $f(x, y, z) = x^2 - y^2$ at a general point of the surface $x^2 + y^2 + z^2 = 4$ in the direction of the outward normal at that point.
(c) $f(x, y, z) = x^2 + y^2 - z^2$ at $(3, 4, 5)$ along the curve of intersection of the two surfaces $2x^2 + 2y^2 - z^2 = 25$ and $x^2 + y^2 = z^2$.
4. (a) Find a vector $V(x, y, z)$ normal to the surface

$$z = \sqrt{x^2 + y^2} + (x^2 + y^2)^{3/2}$$

at a general point (x, y, z) of the surface, $(x, y, z) \neq (0, 0, 0)$.
(b) Find the cosine of the angle 0 between $V(x, y, z)$ and the z-axis and determine the limit of $\cos \theta$ as $(x, y, z) \to (0, 0, 0)$.
5. The two equations $e^u \cos v = x$ and $e^u \sin v = y$ define u and v as functions of x and y, say $u = U(x, y)$ and $v = V(x, y)$. Find explicit formulas for $U(x, y)$ and $V(x, y)$, valid for $x > 0$, and show that the gradient vectors $\nabla U(x, y)$ and $\nabla V(x, y)$ are perpendicular at each point (x, y).
6. Let $f(x, y) = \sqrt{|xy|}$.
(a) Verify that $\partial f/\partial x$ and $\partial f/\partial y$ are both zero at the origin.
(b) Does the surface $z = f(x, y)$ have a tangent plane at the origin? [*Hint:* Consider the section of the surface made by the plane $x = y$.]
7. If (x_0, y_0, z_0) is a point on the surface $z = xy$, then the two lines $z = y_0x$, $y = y_0$ and $z = x_0y$, $x = x_0$ intersect at (x_0, y_0, z_0) and lie on the surface. Verify that the tangent plane to this surface at the point (x_0, y_0, z_0) contains these two lines.
8. Find a Cartesian equation for the tangent plane to the surface $xyz = a^3$ at a general point (x_0, y_0, z_0). Show that the volume of the tetrahedron bounded by this plane and the three coordinate plane is $9a^3/2$.
9. Find a pair of linear Cartesian equations for the line which is tangent to both the surfaces $x^2 + y^2 + 2z^2 = 4$ and $z = e^{x-y}$ at the point $(1, 1, 1)$.
10. Find a constant c such that at any point of intersection of the two spheres

$$(x - c)^2 + y^2 + z^2 = 3 \qquad \text{and} \qquad x^2 + (y - 1)^2 + z^2 = 1$$

the corresponding tangent planes will be perpendicular to each other.
11. If r_1 and r_2 denote the distances from a point (x, y) on an ellipse to its foci, show that the equation $r_1 + r_2$ = constant (satisfied by these distances) implies the relation

$$\boldsymbol{T} \cdot \nabla(r_1 + r_2) = 0,$$

where $\boldsymbol{T}$ is the unit tangent to the curve. Interpret this result geometrically, thereby showing that the tangent makes equal angles with the lines joining (x, y) to the foci.
12. If $\nabla f(x, y, z)$ is always parallel to $x\boldsymbol{i} + y\boldsymbol{j} + z\boldsymbol{k}$, show thatfmust assume equal values at the points (0, 0, a) and (0, 0, -a).

8.18 Derivatives of vector fields

Derivative theory for vector fields is a straightforward extension of that for scalar fields. Let $\boldsymbol{f}: S \to \mathbf{R}^m$ be a vector field defined on a subset S of $\mathbf{R}^n$. If a is an interior point of S and if y is any vector in $\mathbf{R}^n$ we define the derivative $f'(\boldsymbol{a}; y)$ by the formula

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

whenever the limit exists. The derivativef'(a; y) is a vector in $\mathbf{R}^m$.

Let f_k denote the kth component off. We note that the derivative $f'(\boldsymbol{a};\ y)$ exists if and only if $f_k'(\boldsymbol{a};\ y)$ exists for each $k = 1, 2, \ldots, m$, in which case we have

$$f'(\boldsymbol{a};\ y) = (f_1'(\boldsymbol{a};\ y), \ldots, f_m'(\boldsymbol{a};\ y)) = \sum_{k=1}^{m} f_k'(\boldsymbol{a};\ y)\boldsymbol{e}_k,$$

where $\boldsymbol{e}_k$ is the kth unit coordinate vector.

We say that f is *differentiable* at an interior point a if there is a linear transformation

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

such that

$$f(\boldsymbol{a} + \boldsymbol{v}) = f(\boldsymbol{a}) + T_{\boldsymbol{a}}(\boldsymbol{v}) + \|\boldsymbol{v}\|\ E(\boldsymbol{a}, \boldsymbol{v}), \tag{8.16}$$

where $E(\boldsymbol{a}, \text{a}) \to 0$ as $\boldsymbol{v} \to 0$. The first-order Taylor formula (8.16) is to hdd for all $\boldsymbol{v}$ with $\|\boldsymbol{v}\| < r$ for some $r > 0$. The term $E(\boldsymbol{a}, \boldsymbol{v})$ is a vector in $\mathbf{R}^m$. The linear transformation $T_{\boldsymbol{a}}$ is called the *total derivative* off at a.

For scalar fields we proved that $T_{\boldsymbol{a}}(y)$ is the dot product of the gradient vector $\nabla f(\boldsymbol{a})$ withy. For vector fields we will prove that $T_a(y)$ is a vector whose kth component is the dot product $\nabla f_k(\boldsymbol{a}) \cdot y$.

THEOREM 8.9. *Assume f is differentiable at a with total derivative $T_{\boldsymbol{a}}$. Then the derivative $f'(a; y)$ exists for every $\boldsymbol{a}$ in $\mathbf{R}^n$, and we have*

$$T_{\boldsymbol{a}}(y) = f'(\boldsymbol{a}; y). \tag{8.17}$$

Moreover, iff $= (f_1, \ldots, f_m)$ and ify $= (y_1, \ldots, y_n)$, we have

$$T_{\boldsymbol{a}}(y) = \sum_{k=1}^{m} \nabla f_k(\boldsymbol{a}) \cdot y\ \boldsymbol{e}_k = (\nabla f_1(\boldsymbol{a}) \cdot y, \ldots, \nabla f_m(\boldsymbol{a}) \cdot y). \tag{8.18}$$

Proof. We argue as in the scalar case. If $y = 0$, then $f'(\boldsymbol{a}; y) = 0$ and $T_a(O) = 0$. Therefore we can assume that $y \neq 0$. Taking $\boldsymbol{v} = hy$ in the Taylor formula (8.16) we have

$$f(\boldsymbol{a} + hy) - f(\boldsymbol{a}) = T_{\boldsymbol{a}}(hy) + \|hy\|\ E(\boldsymbol{a}, \boldsymbol{v}) = hT_{\boldsymbol{a}}(y) + |h|\ \|y\|\ E(\boldsymbol{a}, \boldsymbol{v}).$$

Dividing by h and letting $h \to 0$ we obtain (8.17).

To prove (8.18) we simply note that

$$f'(\boldsymbol{a}; y) = \sum_{k=1}^{m} f_k'(\boldsymbol{a}; y)\ \boldsymbol{e}_k = \sum_{k=1}^{m} \nabla f_k(\boldsymbol{a}) \cdot y\,\boldsymbol{e}_k.$$

Equation (8.18) can also be written more simply as a matrix product,

$$T_a(y) = Df(a)y,$$

where $Df(a)$ is the $m \times n$ matrix whose kth row is $\nabla f_k(a)$, and where y is regarded as an $n \times 1$ column matrix. The matrix $Df(a)$ is called the *Jacobian matrix* **of** f at a. Its kj entry is the partial derivative $D_j f_k(a)$. Thus, we have

$$Df(a) = \begin{bmatrix} D_1 f_1(a) & D_2 f_1(a) & \cdots & D_n f_1(a) \\ D_1 f_2(a) & D_2 f_2(a) & \cdots & D_n f_2(a) \\ \vdots & & & \vdots \\ D_1 f_m(a) & D_2 f_m(a) & \cdots & D_n f_m(a) \end{bmatrix}$$

The Jacobian matrix $Df(a)$ is defined at each point where the mn partial derivatives $D_j f_k(a)$ exist.

The total derivative T_a is also written as $f'(a)$. The derivative $f'(a)$ is a linear transformation; the Jacobian $Df(a)$ is a matrix representation for this transformation.

The first-order Taylor formula takes the form

$$f(a + v) = f(a) + f'(a)(v) + \|v\| \, E(a, v), \tag{8.19}$$

where $E(a, v) \to 0$ as $v \to 0$. This resembles the one-dimensional Taylor formula. To compute the components of the vector $f'(a)(v)$ we can use the matrix product $Df(a)v$ or formula (8.18) of Theorem 8.9.

8.19 Differentiability implies continuity

THEOREM 8.10. *If a vector field f is differentiable at a, then f is continuous at a.*

Proof. As in the scalar case, we use the Taylor formula to prove this theorem. If we let $v \to 0$ in (8.19) the error term $\|v\| \, E(a, v) \to 0$. The linear part $f'(a)(v)$ also tends to 0 because linear transformations are continuous at 0. This completes the proof.

At this point it is convenient to derive an inequality which will be used in the proof of the chain rule in the next section. The inequality concerns a vector field f differentiable at a; it states that

$$\|f'(a)(v)\| \le M_f(a) \, \|v\|, \qquad \text{where} \quad M_f(a) = \sum_{k=1}^{m} \|\nabla f_k(a)\|. \tag{8.20}$$

To prove this we use Equation (8.18) along with the triangle inequality and the Cauchy-Schwarz inequality to obtain

$$\|f'(a)(v)\| = \left\| \sum_{k=1}^{m} \nabla f_k(a) \cdot v \, e_k \right\| \le \sum_{k=1}^{m} |\nabla f_k(a) \cdot v| \le \sum_{k=1}^{m} \|\nabla f_k(a)\| \, \|v\|.$$

8.20 The chain rule for derivatives of vector fields

THEOREM 8.11. CHAIN RULE. *Let f and g be vector fields such that the composition $h = f \circ g$ is defined in a neighborhood of a point a. Assume that g is differentiable at a, with total derivative $g'(a)$. Let $b = g(a)$ and assume that f is differentiable at b, with total derivative $f'(b)$. Then h is differentiable at a, and the total derivative $h'(a)$ is given by*

$$h'(a) = f'(b) \circ g'(a),$$

the composition of the linear transformations $f'(b)$ and $g'(a)$.

Proof We consider the difference $h(a + y) - h(a)$ for small $\|y\|$, and show that we have a first-order Taylor formula. From the definition of h we have

$$h(a+y) - h(a) = f[g(a+y)] - f[g(a)] = f(b+v) - f(b), \tag{8.21}$$

where $v = g(a + y) - g(a)$. The Taylor formula for $g(a + y)$ gives us

$$v = g'(a)(y) + \|y\|\, E_g(a, y), \qquad \text{where } E_g(a, y) \to 0 \text{ as } y \to 0. \tag{8.22}$$

The Taylor formula for $f(b + v)$ gives us

$$f(b+v) - f(b) = f'(b)(v) + \|v\|\, E_f(b, v), \tag{8.23}$$

where $E_f(b, v) \to 0$ as $v \to 0$. Using (8.22) in (8.23) we obtain

$$\begin{aligned} f(b+v) - f(b) &= f'(b)g'(a)(y) + f'(b)(\|y\|\, E_g(a,y)) + \|v\|\, E_f(b,v) \\ &= f'(b)g'(a)(y) + \|y\|\, E(a, y), \end{aligned} \tag{8.24}$$

where $E(a, O) = 0$ and

$$E(a, y) = f'(b)(E_g(a, y)) + \frac{\|v\|}{\|y\|} E_f(b, v) \qquad \text{if } y \neq O. \tag{8.25}$$

To complete the proof we need to show that $E(a, y) \to 0$ as $y \to 0$.

The first term on the right of (8.25) tends to 0 as $y \to 0$ because $E_g(a, y) \to 0$ as $y \to 0$ and linear transformations are continuous at 0.

In the second term on the right of (8.25) the factor $E_f(b, v) \to 0$ because $v \to 0$ as $y \to 0$. The quotient $\|v\|/\|y\|$ remains bounded because, by (8.22) and (8.20) we have

$$\|v\| \le M_g(a)\, \|y\| + \|y\|\, \|E_g(a, y)\|.$$

Therefore both terms on the right of (8.25) tend to 0 as $y \to O$, so $E(a, y) \to 0$.

Thus, from (8.24) and (8.21) we obtain the Taylor formula

$$h(a + y) - h(a) = f'(b)g'(a)(y) + \|y\| E(a, y),$$

where $E(\boldsymbol{a}, y) \to 0$ as $y \to 0$. This proves that $\boldsymbol{h}$ is differentiable at a and that the total derivative $\boldsymbol{h}'(\boldsymbol{a})$ is equal to the composition $\boldsymbol{f}'(\boldsymbol{b}) \circ$ g'(u).

8.21 Matrix form of the chain rule

Let $\boldsymbol{h} = \boldsymbol{f} \circ$ g , where $\boldsymbol{g}$ is differentiable at a andfis differentiable at $\boldsymbol{b}$ = g(u). The chain rule states that

$$h'(a) = f'(b) \circ g'(u).$$

We can express the chain rule in terms of the Jacobian matrices $D\boldsymbol{h}(\boldsymbol{a})$, $D\boldsymbol{f}(\boldsymbol{b})$, and $D\boldsymbol{g}(\boldsymbol{a})$ which represent the linear transformations h'(a), f'(b), and *g'(u),* respectively. Since composition of linear transformations corresponds to multiplication of their matrices, we obtain

(8.26) $$D\boldsymbol{h}(\boldsymbol{a}) = D\boldsymbol{f}(\boldsymbol{b})\, D\boldsymbol{g}(\boldsymbol{a}), \qquad \text{where } \boldsymbol{b} = \boldsymbol{g}(u).$$

This is called the ***matrix form of the chain rule.*** It can also be written as a set of scalar equations by expressing each matrix in terms of its entries.

Suppose that a $\in \mathbf{R}^p$, 6 = *g(u)* $\in \mathbf{R}^n$, and f(b) $\in \mathbf{R}^m$. Then h(u) $\in \mathbf{R}^m$ and we can write

$$\boldsymbol{g} = (g_1, \ldots, g_n), \qquad \boldsymbol{f} = (f_1, \ldots, f_m), \qquad \boldsymbol{h} = (h_,, \ldots, h_,).$$

Then $D\boldsymbol{h}(\boldsymbol{a})$ is an $m \times p$ matrix, $D\boldsymbol{f}(\boldsymbol{b})$ is an $m \times n$ matrix, and $D\boldsymbol{g}(\boldsymbol{a})$ is an $n \times p$ matrix, given by

$$D\boldsymbol{h}(\boldsymbol{a}) = [D_j h_i(\boldsymbol{a})]_{i,j=1}^{m,p}, \qquad D\boldsymbol{f}(\boldsymbol{b}) = [D_k f_i(\boldsymbol{b})]_{i,k=1}^{m,n}, \qquad D\boldsymbol{g}(\boldsymbol{a}) = [D_j g_k(\boldsymbol{a})]_{k,j=1}^{n,p}.$$

The matrix equation (8.26) is equivalent to mp scalar equations,

$$D_j h_i(\boldsymbol{a}) = \sum_{k=1}^{n} D_k f_i(\boldsymbol{b}) D_j g_k(\boldsymbol{a}), \qquad \text{for } i = 1, 2, \ldots, m \quad \text{and} \quad j = 1, 2, \ldots, p.$$

These equations express the partial derivatives of the components of $\boldsymbol{h}$ in terms of the partial derivatives of the components off and g.

EXAMPLE 1. ***Extended chain rule for scalar fields.*** Suppose f is a scalar field $(m = 1)$. Then h is also a scalar field and there are p equations in the chain rule, one for each of the partial derivatives of h :

$$D_j h(\boldsymbol{a}) = \sum_{k=1}^{n} D_k f(\boldsymbol{b}) D_j g_k(\boldsymbol{a}), \qquad \text{for } j = 1, 2, \ldots, p.$$

The special case $p = 1$ was already considered in Section 8.15. In this case we get only one equation,

$$h'(\boldsymbol{a}) = \sum_{k=1}^{n} D_k f(\boldsymbol{b}) g_k'(\boldsymbol{a}).$$

Now take $p = 2$ and $n = 2$. Write a = (s, t) and $\boldsymbol{b}$ = (x, y) . Then the components x and y are related to s and t by the equations

$$x = g_1(s, t), \qquad y = g_2(s, t).$$

The chain rule gives a pair of equations for the partial derivatives of h:

$$D_1h(s, t) = D_1f(x, y)\, D_1g_1(s, t) + D_2f(x, y)\, D_1g_2(s, t),$$
$$D_2h(s, t) = D_1f(x, y)\, D_2g_1(s, t) + D_2f(x, y)\, D_2g_2(s, t).$$

In the &notation, this pair of equations is usually written as

(8.27) $$\frac{\partial h}{\partial s} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial s}, \qquad \frac{\partial h}{\text{at}} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial t}.$$

EXAMPLE 2. *Polar coordinates.* The temperature of a thin plate is described by a scalar field f, the temperature at (x, y) being $f(x$, y). Polar coordinates x = $r \cos \theta$, $y = r \sin \theta$ are introduced, and the temperature becomes a function of r and θ determined by the equation

$$\varphi(r, \theta) = f(r \cos \theta, r \sin \theta).$$

Express the partial derivatives $\partial\varphi/\partial r$ and $\partial\varphi/\partial\theta$ in terms of the partial derivatives $\partial f/\partial x$ and $\partial f/\partial y$.

Solution. We use the chain rule as expressed in Equation (8.27), writing (r, θ) instead of (s, t), and φ instead of h. The equations

$$X = r \cos \theta, \qquad y = r \sin \theta$$

give us

Jacobian $$\frac{\partial x}{\partial r} = \cos \theta, \quad \frac{\partial y}{\partial r} = \sin \theta, \quad \frac{\partial x}{\partial \theta} = -r \sin \theta, \quad \frac{\partial y}{\partial \theta} = r \cos \theta.$$

Substituting these formulas in (8.27) we obtain

(8.28) $$\frac{\partial \varphi}{\partial r} = \frac{\partial f}{\partial x}\cos \theta + \frac{\partial f}{\partial y}\sin \theta, \qquad \frac{\partial \varphi}{\partial \theta} = -r\frac{\partial f}{\partial x}\sin \theta + r\frac{\partial f}{\partial y}\cos \theta.$$

These are the required formulas for $\partial\varphi/\partial r$ and $\partial\varphi/\partial\theta$.

EXAMPLE 3. *Second-order partial derivatives.* Refer to Example 2 and express the second-order partial derivative $\partial^2\varphi/\partial\theta^2$ in terms of partial derivatives off.

Solution. We begin with the formula for $\partial\varphi/\partial\theta$ in (8.28) and differentiate with respect to θ, treating r as a constant. There are two terms on the right, each of which must be differentiated as a product. Thus we have

$$\text{(8.29)}\quad \begin{aligned}\frac{\partial^2\varphi}{\partial\theta^2} &= -r\frac{\partial f}{\partial x}\frac{\partial(\sin\theta)}{\partial\theta} - r\sin\theta\frac{\partial}{\partial\theta}\left(\frac{\partial f}{\partial x}\right) + r\frac{\partial f}{\partial y}\frac{\partial(\cos\theta)}{\partial\theta} + r\cos\theta\frac{\partial}{\partial\theta}\left(\frac{\partial f}{\partial y}\right)\\ &= -r\cos\theta\frac{\partial f}{\partial x} - r\sin\theta\frac{\partial}{\partial\theta}\left(\frac{\partial f}{\partial x}\right) - r\sin\theta\frac{\partial f}{\partial y} + r\cos\theta\frac{\partial}{\partial\theta}\left(\frac{\partial f}{\partial y}\right).\end{aligned}$$

To compute the derivatives of $\partial f/\partial x$ and $\partial f/\partial y$ with respect to θ we must keep in mind that, as functions of r and θ, $\partial f/\partial x$ and $\partial f/\partial y$ are ***composite functions*** given by

$$\frac{\partial f}{\partial x} = D_1 f(r\cos\theta, r\sin\theta) \qquad\text{and}\qquad \frac{\partial f}{\partial y} = D_2 f(r\cos\theta, r\sin\theta).$$

Therefore, their derivatives with respect to θ must be determined by use of the chain rule. We again use (8.27), with f replaced by $D_1 f$, to obtain

$$\frac{\partial}{\partial\theta}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial(D_1 f)}{\partial x}\frac{\partial x}{\partial\theta} + \frac{\partial(D_1 f)}{\partial y}\frac{\partial y}{\partial\theta} = \frac{\partial^2 f}{\partial x^2}(-r\sin\theta) + \frac{\partial^2 f}{\partial y\,\partial x}(r\cos\theta).$$

Similarly, using (8.27) with f replaced by $D_2 f$, *we* find

$$\frac{\partial}{\partial\theta}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial(D_2 f)}{\partial x}\frac{\partial x}{\partial\theta} + \frac{\partial(D_2 f)}{\partial y}\frac{\partial y}{\partial\theta} = \frac{\partial^2 f}{\partial x\,\partial y}(-r\sin\theta) + \frac{\partial^2 f}{\partial y^2}(r\cos\theta).$$

When these formulas are used in (8.29) we obtain

$$\begin{aligned}\frac{\partial^2\varphi}{\partial\theta^2} = &-r\cos\theta\frac{\partial f}{\partial x} + r^2\sin^2\theta\frac{\partial^2 f}{\partial x^2} - r^2\sin\theta\cos\theta\frac{\partial^2 f}{\partial y\,\partial x}\\ &- r\sin\theta\frac{\partial f}{\partial y} - r^2\sin\theta\cos\theta\frac{\partial^2 f}{\partial x\,\partial y} + r^2\cos^2\theta\frac{\partial^2 f}{\partial y^2}.\end{aligned}$$

This is the required formula for $\partial^2\varphi/\partial\theta^2$. Analogous formulas for the second-order partial derivatives $\partial^2\varphi/\partial r^2$, $\partial^2\varphi/(\partial r\,\partial\theta)$, and $\partial^2\varphi/(\partial\theta\,\partial r)$ are requested in Exercise 5 of the next section.

8.22 Exercises

In these exercises you may assume differentiability of all functions under consideration.

1. The substitution $t = g(x, y)$ converts $F(t)$ into $f(x, y)$, where $f(x, y) = F[g(x, y)]$.
   (a) Show that

$$\frac{\partial f}{\partial x} = F'[g(x, y)]\frac{\partial g}{\partial x} \qquad\text{and}\qquad \frac{\partial f}{\partial y} = F'[g(x, y)]\frac{\partial g}{\partial y}.$$

   (b) Consider the special case $F(t) = e^{\sin t}$, $g(x, y) = \cos(x^2 + y^2)$. Compute $\partial f/\partial x$ and $\partial f/\partial y$ by use of the formulas in part (a). To check your result, determine $f(x, y)$ explicitly in terms of x and y and compute $\partial f/\partial x$ and $\partial f/\partial y$ directly from f.