14.5 Directional Derivatives and Gradient Vectors

If you look at the map (Figure 14.23) showing contours on the West Point Area along the Hudson River in New York, you will notice that the tributary streams flow perpendicular to the contours. The streams are following paths of steepest descent so the waters reach the Hudson as quickly as possible. Therefore, the instantaneous rate of change in a stream's altitude above sea level has a particular direction. In this section, you see why this direction, called the "downhill" direction, is perpendicular to the contours.

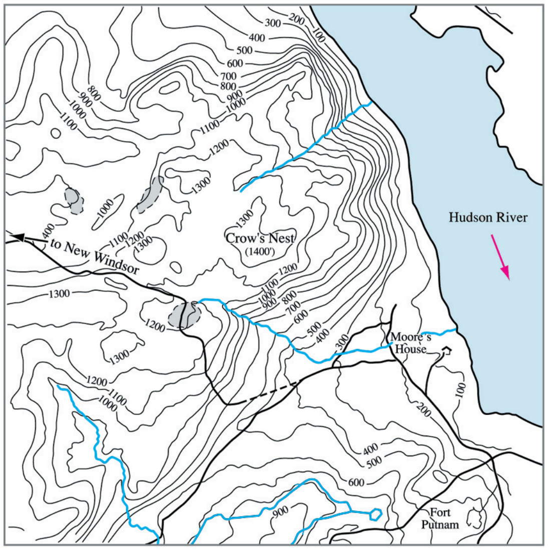

FIGURE 14.23 Contours of the West Point Area in New York show streams, which follow paths of steepest descent, running perpendicular to the contours.

Directional Derivatives in the Plane

We know from Section 14.4 that if $f(x, y)$ is differentiable, then the rate at which f changes with respect to t along a differentiable curve $x = g(t), y = h(t)$ is

$$\frac{df}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt}.$$

At any point $P_0(x_0, y_0) = P_0(g(t_0), h(t_0))$, this equation gives the rate of change of f with respect to increasing t and therefore depends, among other things, on the direction of motion along the curve. If the curve is a straight line and t is the arc length parameter along the line measured from P_0 in the direction of a given unit vector $\mathbf{u}$, then df/dt is the rate of change of f with respect to distance in its domain in the direction of $\mathbf{u}$. By varying $\mathbf{u}$, we find the rates at which f changes with respect to distance as we move through P_0 in different directions. We now define this idea more precisely.

Suppose that the function $f(x, y)$ is defined throughout a region R in the xy-plane, that $P_0(x_0, y_0)$ is a point in R, and that $\mathbf{u} = u_1\mathbf{i} + u_2\mathbf{j}$ is a unit vector. Then the equations

$$x = x_0 + su_1, \qquad y = y_0 + su_2$$

parametrize the line through P_0 parallel to $\mathbf{u}$. If the parameter s measures arc length from P_0 in the direction of $\mathbf{u}$, we find the rate of change of f at P_0 in the direction of $\mathbf{u}$ by calculating df/ds at P_0 (Figure 14.24).

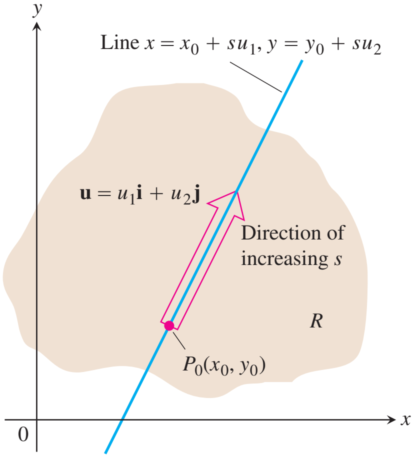

FIGURE 14.24 The rate of change of f in the direction of $\mathbf{u}$ at a point P_0 is the rate at which f changes along this line at P_0.

DEFINITION Directional Derivative

The **derivative of f at $P_0(x_0, y_0)$ in the direction of the unit vector $\mathbf{u} = u_1\mathbf{i} + u_2\mathbf{j}$** is the number

$$\left(\frac{df}{ds}\right)_{\mathbf{u},P_0} = \lim_{s\to 0}\frac{f(x_0+su_1, y_0+su_2) - f(x_0,y_0)}{s}, \tag{1}$$

provided the limit exists.

The directional derivative is also denoted by

$$(D_{\mathbf{u}}f)_{P_0}.$$

"The derivative of f at P_0 in the direction of $\mathbf{u}$"

EXAMPLE 1 Finding a Directional Derivative Using the Definition

Find the derivative of

$$f(x,y) = x^2 + xy$$

at $P_0(1, 2)$ in the direction of the unit vector $\mathbf{u} = \left(1/\sqrt{2}\right)\mathbf{i} + \left(1/\sqrt{2}\right)\mathbf{j}$.

Solution

$$\left(\frac{df}{ds}\right)_{\mathbf{u},P_0} = \lim_{s\to 0}\frac{f(x_0+su_1, y_0+su_2) - f(x_0,y_0)}{s} \qquad \text{Equation (1)}$$

$$= \lim_{s\to 0}\frac{f\left(1+s\cdot\frac{1}{\sqrt{2}}, 2+s\cdot\frac{1}{\sqrt{2}}\right) - f(1,2)}{s}$$

$$= \lim_{s\to 0}\frac{\left(1+\frac{s}{\sqrt{2}}\right)^2 + \left(1+\frac{s}{\sqrt{2}}\right)\left(2+\frac{s}{\sqrt{2}}\right) - (1^2+1\cdot 2)}{s}$$

$$= \lim_{s\to 0}\frac{\left(1+\frac{2s}{\sqrt{2}}+\frac{s^2}{2}\right) + \left(2+\frac{3s}{\sqrt{2}}+\frac{s^2}{2}\right) - 3}{s}$$

$$= \lim_{s\to 0}\frac{\frac{5s}{\sqrt{2}}+s^2}{s} = \lim_{s\to 0}\left(\frac{5}{\sqrt{2}}+s\right) = \left(\frac{5}{\sqrt{2}}+0\right) = \frac{5}{\sqrt{2}}.$$

The rate of change of $f(x,y) = x^2 + xy$ at $P_0(1, 2)$ in the direction $\mathbf{u} = \left(1/\sqrt{2}\right)\mathbf{i} + \left(1/\sqrt{2}\right)\mathbf{j}$ is $5/\sqrt{2}$. ■

Interpretation of the Directional Derivative

The equation $z = f(x, y)$ represents a surface S in space. If $z_0 = f(x_0, y_0)$, then the point $P(x_0, y_0, z_0)$ lies on S. The vertical plane that passes through P and $P_0(x_0, y_0)$ parallel to $\mathbf{u}$

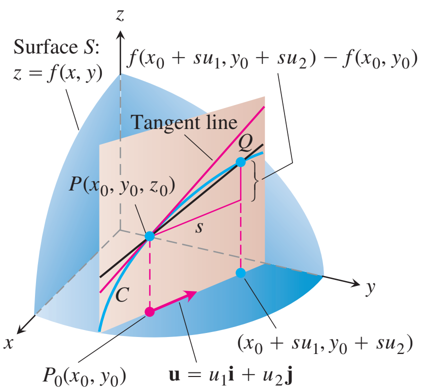

FIGURE 14.25 The slope of curve C at P_0 is $\lim_{Q \to P}$ slope (PQ); this is the directional derivative

$$\left(\frac{df}{ds}\right)_{\mathbf{u},P_0} = (D_{\mathbf{u}}f)_{P_0}.$$

intersects S in a curve C (Figure 14.25). The rate of change of f in the direction of $\mathbf{u}$ is the slope of the tangent to C at P.

When $\mathbf{u} = \mathbf{i}$, the directional derivative at P_0 is $\partial f/\partial x$ evaluated at (x_0, y_0). When $\mathbf{u} = \mathbf{j}$, the directional derivative at P_0 is $\partial f/\partial y$ evaluated at (x_0, y_0). The directional derivative generalizes the two partial derivatives. We can now ask for the rate of change of f in any direction $\mathbf{u}$, not just the directions $\mathbf{i}$ and $\mathbf{j}$.

Here's a physical interpretation of the directional derivative. Suppose that $T = f(x, y)$ is the temperature at each point (x, y) over a region in the plane. Then $f(x_0, y_0)$ is the temperature at the point $P_0(x_0, y_0)$ and $(D_{\mathbf{u}}f)_{P_0}$ is the instantaneous rate of change of the temperature at P_0 stepping off in the direction $\mathbf{u}$.

Calculation and Gradients

We now develop an efficient formula to calculate the directional derivative for a differentiable function f. We begin with the line

$$x = x_0 + su_1, \qquad y = y_0 + su_2, \tag{2}$$

through $P_0(x_0, y_0)$, parametrized with the arc length parameter s increasing in the direction of the unit vector $\mathbf{u} = u_1\mathbf{i} + u_2\mathbf{j}$. Then

$$\left(\frac{df}{ds}\right)_{\mathbf{u},P_0} = \left(\frac{\partial f}{\partial x}\right)_{P_0}\frac{dx}{ds} + \left(\frac{\partial f}{\partial y}\right)_{P_0}\frac{dy}{ds} \qquad \text{Chain Rule for differentiable } f$$

$$= \left(\frac{\partial f}{\partial x}\right)_{P_0}\cdot u_1 + \left(\frac{\partial f}{\partial y}\right)_{P_0}\cdot u_2 \qquad \text{From Equations (2), } dx/ds = u_1 \text{ and } dy/ds = u_2$$

$$= \underbrace{\left[\left(\frac{\partial f}{\partial x}\right)_{P_0}\mathbf{i} + \left(\frac{\partial f}{\partial y}\right)_{P_0}\mathbf{j}\right]}_{\text{Gradient of } f \text{ at } P_0}\cdot\underbrace{\left[u_1\mathbf{i} + u_2\mathbf{j}\right]}_{\text{Direction } \mathbf{u}}. \tag{3}$$

You Try It

> **DEFINITION Gradient Vector**
> The **gradient vector (gradient)** of $f(x, y)$ at a point $P_0(x_0, y_0)$ is the vector
> $$\nabla f = \frac{\partial f}{\partial x}\mathbf{i} + \frac{\partial f}{\partial y}\mathbf{j}$$
> obtained by evaluating the partial derivatives of f at P_0.

The notation ∇f is read "grad f" as well as "gradient of f" and "del f." The symbol ∇ by itself is read "del." Another notation for the gradient is grad f, read the way it is written.

Equation (3) says that the derivative of a differentiable function f in the direction of $\mathbf{u}$ at P_0 is the dot product of $\mathbf{u}$ with the gradient of f at P_0.

THEOREM 9 The Directional Derivative Is a Dot Product

If $f(x, y)$ is differentiable in an open region containing $P_0(x_0, y_0)$, then

$$\left(\frac{df}{ds}\right)_{\mathbf{u},P_0} = (\nabla f)_{P_0} \cdot \mathbf{u}, \tag{4}$$

the dot product of the gradient f at P_0 and $\mathbf{u}$.

Animation Video You Try It

EXAMPLE 2 Finding the Directional Derivative Using the Gradient

Find the derivative of $f(x, y) = xe^y + \cos(xy)$ at the point (2, 0) in the direction of $\mathbf{v} = 3\mathbf{i} - 4\mathbf{j}$.

Solution The direction of $\mathbf{v}$ is the unit vector obtained by dividing $\mathbf{v}$ by its length:

$$\mathbf{u} = \frac{\mathbf{v}}{|\mathbf{v}|} = \frac{\mathbf{v}}{5} = \frac{3}{5}\mathbf{i} - \frac{4}{5}\mathbf{j}.$$

The partial derivatives of f are everywhere continuous and at (2, 0) are given by

$$f_x(2, 0) = (e^y - y\sin(xy))_{(2,0)} = e^0 - 0 = 1$$
$$f_y(2, 0) = (xe^y - x\sin(xy))_{(2,0)} = 2e^0 - 2 \cdot 0 = 2.$$

The gradient of f at (2, 0) is

$$\nabla f|_{(2,0)} = f_x(2, 0)\mathbf{i} + f_y(2, 0)\mathbf{j} = \mathbf{i} + 2\mathbf{j}$$

(Figure 14.26). The derivative of f at (2, 0) in the direction of $\mathbf{v}$ is therefore

$$(D_{\mathbf{u}}f)|_{(2,0)} = \nabla f|_{(2,0)} \cdot \mathbf{u} \qquad \text{Equation (4)}$$
$$= (\mathbf{i} + 2\mathbf{j}) \cdot \left(\frac{3}{5}\mathbf{i} - \frac{4}{5}\mathbf{j}\right) = \frac{3}{5} - \frac{8}{5} = -1.$$

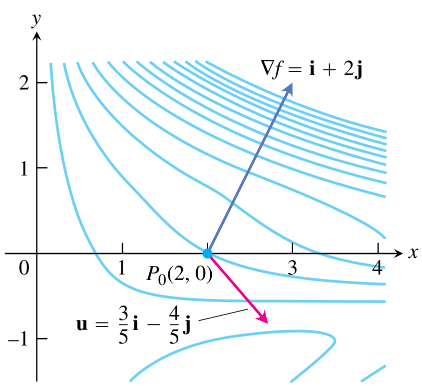

FIGURE 14.26 Picture ∇f as a vector in the domain of f. In the case of $f(x, y) = xe^y + \cos(xy)$, the domain is the entire plane. The rate at which f changes at (2, 0) in the direction $\mathbf{u} = (3/5)\mathbf{i} - (4/5)\mathbf{j}$ is $\nabla f \cdot \mathbf{u} = -1$ (Example 2).

Evaluating the dot product in the formula

$$D_{\mathbf{u}}f = \nabla f \cdot \mathbf{u} = |\nabla f||\mathbf{u}|\cos\theta = |\nabla f|\cos\theta,$$

where θ is the angle between the vectors $\mathbf{u}$ and ∇f, reveals the following properties.

Properties of the Directional Derivative $D_{\mathbf{u}}f = \nabla f \cdot \mathbf{u} = |\nabla f|\cos\theta$

1. The function f increases most rapidly when $\cos\theta = 1$ or when $\mathbf{u}$ is the direction of ∇f. That is, at each point P in its domain, f increases most rapidly in the direction of the gradient vector ∇f at P. The derivative in this direction is
$$D_{\mathbf{u}}f = |\nabla f|\cos(0) = |\nabla f|.$$
2. Similarly, f decreases most rapidly in the direction of $-\nabla f$. The derivative in this direction is $D_{\mathbf{u}}f = |\nabla f|\cos(\pi) = -|\nabla f|$.
3. Any direction $\mathbf{u}$ orthogonal to a gradient $\nabla f \neq 0$ is a direction of zero change in f because θ then equals $\pi/2$ and
$$D_{\mathbf{u}}f = |\nabla f|\cos(\pi/2) = |\nabla f| \cdot 0 = 0.$$

As we discuss later, these properties hold in three dimensions as well as two.

EXAMPLE 3 Finding Directions of Maximal, Minimal, and Zero Change

Find the directions in which $f(x, y) = (x^2/2) + (y^2/2)$

(a) Increases most rapidly at the point (1, 1)

(b) Decreases most rapidly at (1, 1).

(c) What are the directions of zero change in f at (1, 1)?

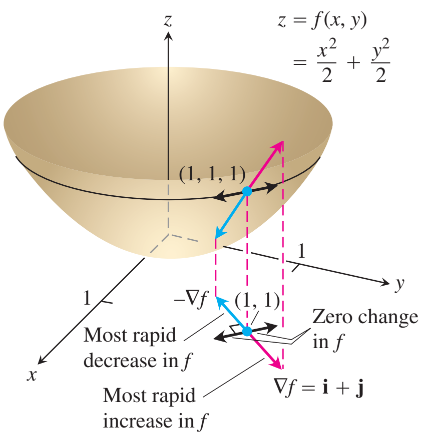

FIGURE 14.27 The direction in which $f(x, y) = (x^2/2) + (y^2/2)$ increases most rapidly at (1, 1) is the direction of $\nabla f|_{(1,1)} = \mathbf{i} + \mathbf{j}$. It corresponds to the direction of steepest ascent on the surface at (1, 1, 1) (Example 3).

Solution

(a) The function increases most rapidly in the direction of ∇f at (1, 1). The gradient there is

$$(\nabla f)_{(1,1)} = (x\mathbf{i} + y\mathbf{j})_{(1,1)} = \mathbf{i} + \mathbf{j}.$$

Its direction is

$$\mathbf{u} = \frac{\mathbf{i} + \mathbf{j}}{|\mathbf{i} + \mathbf{j}|} = \frac{\mathbf{i} + \mathbf{j}}{\sqrt{(1)^2 + (1)^2}} = \frac{1}{\sqrt{2}}\mathbf{i} + \frac{1}{\sqrt{2}}\mathbf{j}.$$

(b) The function decreases most rapidly in the direction of $-\nabla f$ at (1, 1), which is

$$-\mathbf{u} = -\frac{1}{\sqrt{2}}\mathbf{i} - \frac{1}{\sqrt{2}}\mathbf{j}.$$

(c) The directions of zero change at (1, 1) are the directions orthogonal to ∇f:

$$\mathbf{n} = -\frac{1}{\sqrt{2}}\mathbf{i} + \frac{1}{\sqrt{2}}\mathbf{j} \quad \text{and} \quad -\mathbf{n} = \frac{1}{\sqrt{2}}\mathbf{i} - \frac{1}{\sqrt{2}}\mathbf{j}.$$

See Figure 14.27. ■

Gradients and Tangents to Level Curves

If a differentiable function $f(x, y)$ has a constant value c along a smooth curve $\mathbf{r} = g(t)\mathbf{i} + h(t)\mathbf{j}$ (making the curve a level curve of f), then $f(g(t), h(t)) = c$. Differentiating both sides of this equation with respect to t leads to the equations

$$\frac{d}{dt} f(g(t), h(t)) = \frac{d}{dt}(c)$$

$$\frac{\partial f}{\partial x}\frac{dg}{dt} + \frac{\partial f}{\partial y}\frac{dh}{dt} = 0 \qquad \text{Chain Rule}$$

$$\underbrace{\left(\frac{\partial f}{\partial x}\mathbf{i} + \frac{\partial f}{\partial y}\mathbf{j}\right)}_{\nabla f} \cdot \underbrace{\left(\frac{dg}{dt}\mathbf{i} + \frac{dh}{dt}\mathbf{j}\right)}_{\frac{d\mathbf{r}}{dt}} = 0. \qquad (5)$$

Equation (5) says that ∇f is normal to the tangent vector $d\mathbf{r}/dt$, so it is normal to the curve.

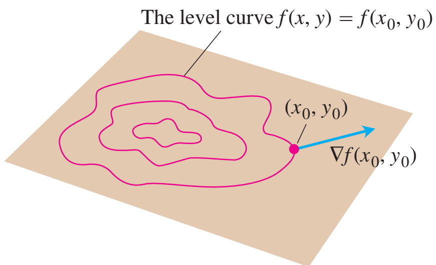

FIGURE 14.28 The gradient of a differentiable function of two variables at a point is always normal to the function's level curve through that point.

At every point (x_0, y_0) in the domain of a differentiable function $f(x, y)$, the gradient of f is normal to the level curve through (x_0, y_0) (Figure 14.28).

Equation (5) validates our observation that streams flow perpendicular to the contours in topographical maps (see Figure 14.23). Since the downflowing stream will reach its destination in the fastest way, it must flow in the direction of the negative gradient vectors from Property 2 for the directional derivative. Equation (5) tells us these directions are perpendicular to the level curves.

This observation also enables us to find equations for tangent lines to level curves. They are the lines normal to the gradients. The line through a point $P_0(x_0, y_0)$ normal to a vector $\mathbf{N} = A\mathbf{i} + B\mathbf{j}$ has the equation

$$A(x - x_0) + B(y - y_0) = 0$$

(Exercise 35). If $\mathbf{N}$ is the gradient $(\nabla f)_{(x_0, y_0)} = f_x(x_0, y_0)\mathbf{i} + f_y(x_0, y_0)\mathbf{j}$, the equation is the tangent line given by

$$f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0) = 0. \qquad (6)$$

EXAMPLE 4 Finding the Tangent Line to an Ellipse

Find an equation for the tangent to the ellipse

$$\frac{x^2}{4} + y^2 = 2$$

(Figure 14.29) at the point $(-2, 1)$.

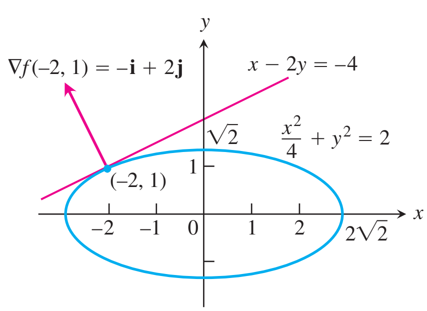

FIGURE 14.29 We can find the tangent to the ellipse $(x^2/4) + y^2 = 2$ by treating the ellipse as a level curve of the function $f(x, y) = (x^2/4) + y^2$ (Example 4).

Solution The ellipse is a level curve of the function

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

The gradient of f at $(-2, 1)$ is

$$\nabla f|_{(-2,1)} = \left(\frac{x}{2}\mathbf{i} + 2y\mathbf{j}\right)_{(-2,1)} = -\mathbf{i} + 2\mathbf{j}.$$

The tangent is the line

$$(-1)(x + 2) + (2)(y - 1) = 0 \qquad \text{Equation (6)}$$
$$x - 2y = -4.$$

■

If we know the gradients of two functions f and g, we automatically know the gradients of their constant multiples, sum, difference, product, and quotient. You are asked to establish the following rules in Exercise 36. Notice that these rules have the same form as the corresponding rules for derivatives of single-variable functions.

Algebra Rules for Gradients

1. *Constant Multiple Rule:* $\nabla(kf) = k\nabla f$ (any number k)
2. *Sum Rule:* $\nabla(f + g) = \nabla f + \nabla g$
3. *Difference Rule:* $\nabla(f - g) = \nabla f - \nabla g$
4. *Product Rule:* $\nabla(fg) = f\nabla g + g\nabla f$
5. *Quotient Rule:* $\nabla\left(\frac{f}{g}\right) = \frac{g\nabla f - f\nabla g}{g^2}$

EXAMPLE 5 Illustrating the Gradient Rules

We illustrate the rules with

$$\begin{aligned} f(x, y) &= x - y \qquad & g(x, y) &= 3y \\ \nabla f &= \mathbf{i} - \mathbf{j} & \nabla g &= 3\mathbf{j}. \end{aligned}$$

We have

1. $\nabla(2f) = \nabla(2x - 2y) = 2\mathbf{i} - 2\mathbf{j} = 2\nabla f$
2. $\nabla(f + g) = \nabla(x + 2y) = \mathbf{i} + 2\mathbf{j} = \nabla f + \nabla g$
3. $\nabla(f - g) = \nabla(x - 4y) = \mathbf{i} - 4\mathbf{j} = \nabla f - \nabla g$
4. $$\begin{aligned} \nabla(fg) &= \nabla(3xy - 3y^2) = 3y\mathbf{i} + (3x - 6y)\mathbf{j} \\ &= 3y(\mathbf{i} - \mathbf{j}) + 3y\mathbf{j} + (3x - 6y)\mathbf{j} \\ &= 3y(\mathbf{i} - \mathbf{j}) + (3x - 3y)\mathbf{j} \\ &= 3y(\mathbf{i} - \mathbf{j}) + (x - y)3\mathbf{j} = g\nabla f + f\nabla g \end{aligned}$$
5. $$\begin{aligned} \nabla\left(\frac{f}{g}\right) &= \nabla\left(\frac{x - y}{3y}\right) = \nabla\left(\frac{x}{3y} - \frac{1}{3}\right) \\ &= \frac{1}{3y}\mathbf{i} - \frac{x}{3y^2}\mathbf{j} \\ &= \frac{3y\mathbf{i} - 3x\mathbf{j}}{9y^2} = \frac{3y(\mathbf{i} - \mathbf{j}) - (3x - 3y)\mathbf{j}}{9y^2} \\ &= \frac{3y(\mathbf{i} - \mathbf{j}) - (x - y)3\mathbf{j}}{9y^2} = \frac{g\nabla f - f\nabla g}{g^2}. \end{aligned}$$

■

Functions of Three Variables

For a differentiable function $f(x, y, z)$ and a unit vector $\mathbf{u} = u_1\mathbf{i} + u_2\mathbf{j} + u_3\mathbf{k}$ in space, we have

$$\nabla f = \frac{\partial f}{\partial x}\mathbf{i} + \frac{\partial f}{\partial y}\mathbf{j} + \frac{\partial f}{\partial z}\mathbf{k}$$

and

$$D_{\mathbf{u}}f = \nabla f \cdot \mathbf{u} = \frac{\partial f}{\partial x}u_1 + \frac{\partial f}{\partial y}u_2 + \frac{\partial f}{\partial z}u_3.$$

The directional derivative can once again be written in the form

$$D_{\mathbf{u}}f = \nabla f \cdot \mathbf{u} = |\nabla f||u|\cos\theta = |\nabla f|\cos\theta,$$

so the properties listed earlier for functions of two variables continue to hold. At any given point, f increases most rapidly in the direction of ∇f and decreases most rapidly in the direction of $-\nabla f$. In any direction orthogonal to ∇f, the derivative is zero.

EXAMPLE 6 Finding Directions of Maximal, Minimal, and Zero Change

(a) Find the derivative of $f(x, y, z) = x^3 - xy^2 - z$ at $P_0(1, 1, 0)$ in the direction of $\mathbf{v} = 2\mathbf{i} - 3\mathbf{j} + 6\mathbf{k}$.

(b) In what directions does f change most rapidly at P_0, and what are the rates of change in these directions?

Solution

(a) The direction of $\mathbf{v}$ is obtained by dividing $\mathbf{v}$ by its length:

$$|\mathbf{v}| = \sqrt{(2)^2 + (-3)^2 + (6)^2} = \sqrt{49} = 7$$

$$\mathbf{u} = \frac{\mathbf{v}}{|\mathbf{v}|} = \frac{2}{7}\mathbf{i} - \frac{3}{7}\mathbf{j} + \frac{6}{7}\mathbf{k}.$$

The partial derivatives of f at P_0 are

$$f_x = (3x^2 - y^2)_{(1,1,0)} = 2, \qquad f_y = -2xy|_{(1,1,0)} = -2, \qquad f_z = -1|_{(1,1,0)} = -1.$$

The gradient of f at P_0 is

$$\nabla f|_{(1,1,0)} = 2\mathbf{i} - 2\mathbf{j} - \mathbf{k}.$$

The derivative of f at P_0 in the direction of $\mathbf{v}$ is therefore

$$(D_{\mathbf{u}}f)_{(1,1,0)} = \nabla f|_{(1,1,0)} \cdot \mathbf{u} = (2\mathbf{i} - 2\mathbf{j} - \mathbf{k}) \cdot \left(\frac{2}{7}\mathbf{i} - \frac{3}{7}\mathbf{j} + \frac{6}{7}\mathbf{k}\right)$$

$$= \frac{4}{7} + \frac{6}{7} - \frac{6}{7} = \frac{4}{7}.$$

(b) The function increases most rapidly in the direction of $\nabla f = 2\mathbf{i} - 2\mathbf{j} - \mathbf{k}$ and decreases most rapidly in the direction of $-\nabla f$. The rates of change in the directions are, respectively,

$$|\nabla f| = \sqrt{(2)^2 + (-2)^2 + (-1)^2} = \sqrt{9} = 3 \quad \text{and} \quad -|\nabla f| = -3.$$

■