

# Lecture 2

Tangent Plane. The Chain Rule.  
The Directional derivatives and  
the Gradient.

## Tangent Planes and Linear Approximations

One of the most important ideas in single-variable calculus is that as we zoom in toward a point on the graph of a differentiable function, the graph becomes indistinguishable from its tangent line and we can approximate the function by a linear function. (See Section 3.10.) Here we develop similar ideas in three dimensions. As we zoom in toward a point on a surface that is the graph of a differentiable function of two variables, the surface looks more and more like a plane (its tangent plane) and we can approximate the function by a linear function of two variables. We also extend the idea of a differential to functions of two or more variables.

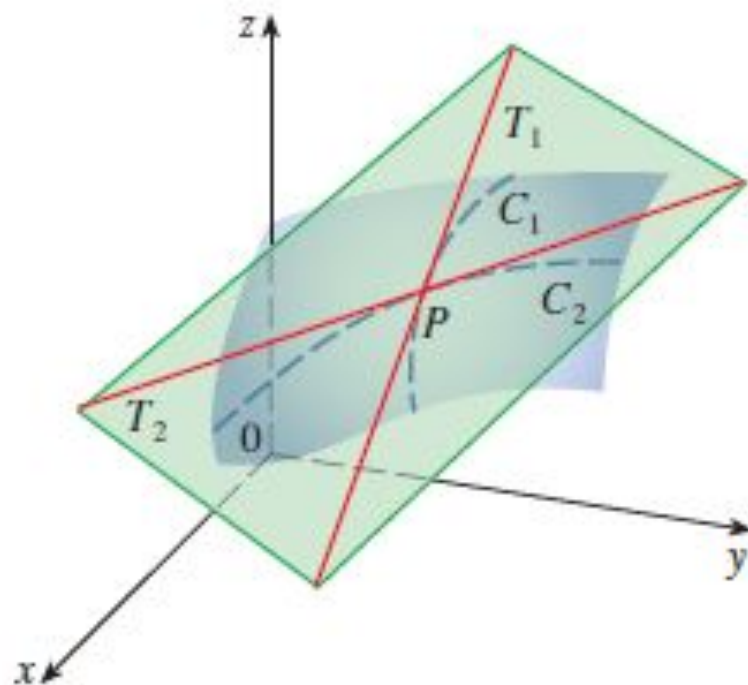
### ■ Tangent Planes

Suppose a surface  $S$  has equation  $z = f(x, y)$ , where  $f$  has continuous first partial derivatives, and let  $P(x_0, y_0, z_0)$  be a point on  $S$ . As in Section 14.3, let  $C_1$  and  $C_2$  be the curves obtained by intersecting the vertical planes  $y = y_0$  and  $x = x_0$  with the surface  $S$ . Then the point  $P$  lies on both  $C_1$  and  $C_2$ . Let  $T_1$  and  $T_2$  be the tangent lines to the curves  $C_1$  and  $C_2$  at the point  $P$ . Then the **tangent plane** to the surface  $S$  at the point  $P$  is defined to be the plane that contains both tangent lines  $T_1$  and  $T_2$ . (See Figure 1.)

We will see in Section 14.6 that if  $C$  is any other curve that lies on the surface  $S$  and passes through  $P$ , then its tangent line at  $P$  also lies in the tangent plane. Therefore you can think of the tangent plane to  $S$  at  $P$  as consisting of all possible tangent lines at  $P$  to curves that lie on  $S$  and pass through  $P$ . The tangent plane at  $P$  is the plane that most closely approximates the surface  $S$  near the point  $P$ .

We know from Equation 12.5.7 that any plane passing through the point  $P(x_0, y_0, z_0)$  has an equation of the form

$$A(x - x_0) + B(y - y_0) + C(z - z_0) = 0$$



**FIGURE 1**

The tangent plane contains the tangent lines  $T_1$  and  $T_2$ .

By dividing this equation by  $C$  and letting  $a = -A/C$  and  $b = -B/C$ , we can write it in the form

$$\boxed{1} \quad z - z_0 = a(x - x_0) + b(y - y_0)$$

If Equation 1 represents the tangent plane at  $P$ , then its intersection with the plane  $y = y_0$  must be the tangent line  $T_1$ . Setting  $y = y_0$  in Equation 1 gives

$$z - z_0 = a(x - x_0) \quad \text{where } y = y_0$$

and we recognize this as the equation (in point-slope form) of a line with slope  $a$ . But from Section 14.3 we know that the slope of the tangent  $T_1$  is  $f_x(x_0, y_0)$ . Therefore  $a = f_x(x_0, y_0)$ .

Similarly, putting  $x = x_0$  in Equation 1, we get  $z - z_0 = b(y - y_0)$ , which must represent the tangent line  $T_2$ , so  $b = f_y(x_0, y_0)$ .

**2 Equation of a Tangent Plane** Suppose  $f$  has continuous partial derivatives. An equation of the tangent plane to the surface  $z = f(x, y)$  at the point  $P(x_0, y_0, z_0)$  is

$$z - z_0 = f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0)$$

**EXAMPLE 1** Find the tangent plane to the elliptic paraboloid  $z = 2x^2 + y^2$  at the point  $(1, 1, 3)$ .

**SOLUTION** Let  $f(x, y) = 2x^2 + y^2$ . Then

$$f_x(x, y) = 4x \quad f_y(x, y) = 2y$$

$$f_x(1, 1) = 4 \quad f_y(1, 1) = 2$$

Then (2) gives the equation of the tangent plane at  $(1, 1, 3)$  as

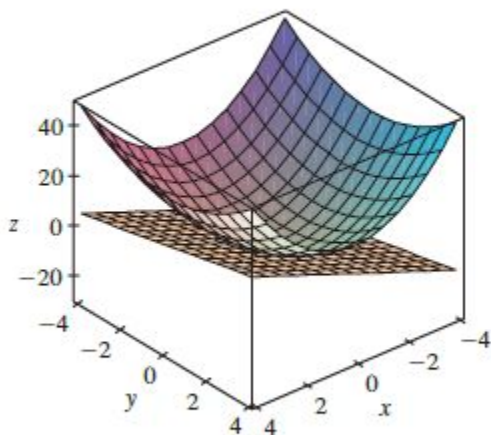
$$z - 3 = 4(x - 1) + 2(y - 1)$$

or

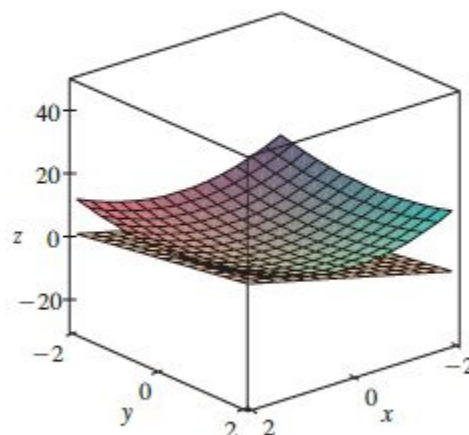
$$z = 4x + 2y - 3$$



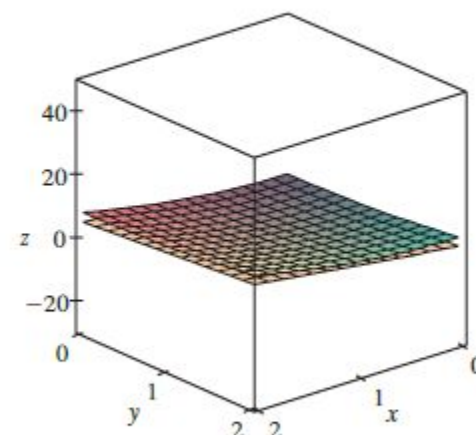
Figure 2(a) shows the elliptic paraboloid and its tangent plane at  $(1, 1, 3)$  that we found in Example 1. In parts (b) and (c) we zoom in toward the point  $(1, 1, 3)$ . Notice that the more we zoom in, the flatter the graph appears and the more it resembles its tangent plane.



(a)



(b)



(c)

**FIGURE 2** The elliptic paraboloid  $z = 2x^2 + y^2$  appears to coincide with its tangent plane as we zoom in toward  $(1, 1, 3)$ .

## The Chain Rule

Recall that the Chain Rule for functions of a single variable gives the rule for differentiating a composite function: if  $y = f(x)$  and  $x = g(t)$ , where  $f$  and  $g$  are differentiable functions, then  $y$  is indirectly a differentiable function of  $t$  and

$$\frac{dy}{dt} = \frac{dy}{dx} \frac{dx}{dt}$$

In this section we extend the Chain Rule to functions of more than one variable.

### ■ The Chain Rule: Case 1

For functions of more than one variable, the Chain Rule has several versions, each of them giving a rule for differentiating a composite function. The first version (Theorem 1) deals with the case where  $z = f(x, y)$  and each of the variables  $x$  and  $y$  is, in turn, a function of a variable  $t$ . This means that  $z$  is indirectly a function of  $t$ ,  $z = f(g(t), h(t))$ , and the Chain Rule gives a formula for differentiating  $z$  as a function of  $t$ . We assume that  $f$  is differentiable (Definition 14.4.7). Recall that this is the case when  $f_x$  and  $f_y$  are continuous (Theorem 14.4.8).

**1 The Chain Rule (Case 1)** Suppose that  $z = f(x, y)$  is a differentiable function of  $x$  and  $y$ , where  $x = g(t)$  and  $y = h(t)$  are both differentiable functions of  $t$ . Then  $z$  is a differentiable function of  $t$  and

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

**PROOF** A change of  $\Delta t$  in  $t$  produces changes of  $\Delta x$  in  $x$  and  $\Delta y$  in  $y$ . These, in turn, produce a change of  $\Delta z$  in  $z$ , and from Definition 14.4.7 we have

$$\Delta z = \frac{\partial f}{\partial x} \Delta x + \frac{\partial f}{\partial y} \Delta y + \varepsilon_1 \Delta x + \varepsilon_2 \Delta y$$

where  $\varepsilon_1 \rightarrow 0$  and  $\varepsilon_2 \rightarrow 0$  as  $(\Delta x, \Delta y) \rightarrow (0, 0)$ . [If the functions  $\varepsilon_1$  and  $\varepsilon_2$  are not defined at  $(0, 0)$ , we can define them to be 0 there.] Dividing both sides of this equation by  $\Delta t$ , we have

$$\frac{\Delta z}{\Delta t} = \frac{\partial f}{\partial x} \frac{\Delta x}{\Delta t} + \frac{\partial f}{\partial y} \frac{\Delta y}{\Delta t} + \varepsilon_1 \frac{\Delta x}{\Delta t} + \varepsilon_2 \frac{\Delta y}{\Delta t}$$

If we now let  $\Delta t \rightarrow 0$ , then  $\Delta x = g(t + \Delta t) - g(t) \rightarrow 0$  because  $g$  is differentiable and therefore continuous. Similarly,  $\Delta y \rightarrow 0$ . This, in turn, means that  $\varepsilon_1 \rightarrow 0$  and  $\varepsilon_2 \rightarrow 0$ , so

$$\begin{aligned} \frac{dz}{dt} &= \lim_{\Delta t \rightarrow 0} \frac{\Delta z}{\Delta t} \\ &= \frac{\partial f}{\partial x} \lim_{\Delta t \rightarrow 0} \frac{\Delta x}{\Delta t} + \frac{\partial f}{\partial y} \lim_{\Delta t \rightarrow 0} \frac{\Delta y}{\Delta t} + \left( \lim_{\Delta t \rightarrow 0} \varepsilon_1 \right) \lim_{\Delta t \rightarrow 0} \frac{\Delta x}{\Delta t} + \left( \lim_{\Delta t \rightarrow 0} \varepsilon_2 \right) \lim_{\Delta t \rightarrow 0} \frac{\Delta y}{\Delta t} \\ &= \frac{\partial f}{\partial x} \frac{dx}{dt} + \frac{\partial f}{\partial y} \frac{dy}{dt} + 0 \cdot \frac{dx}{dt} + 0 \cdot \frac{dy}{dt} \\ &= \frac{\partial f}{\partial x} \frac{dx}{dt} + \frac{\partial f}{\partial y} \frac{dy}{dt} \end{aligned}$$



**EXAMPLE 1** If  $z = x^2y + 3xy^4$ , where  $x = \sin 2t$  and  $y = \cos t$ , find  $dz/dt$  when  $t = 0$ .

**SOLUTION** The Chain Rule gives

$$\begin{aligned}\frac{dz}{dt} &= \frac{\partial z}{\partial x} \frac{dx}{dt} + \frac{\partial z}{\partial y} \frac{dy}{dt} \\ &= (2xy + 3y^4)(2 \cos 2t) + (x^2 + 12xy^3)(-\sin t)\end{aligned}$$

It's not necessary to substitute the expressions for  $x$  and  $y$  in terms of  $t$ . We simply observe that when  $t = 0$ , we have  $x = \sin 0 = 0$  and  $y = \cos 0 = 1$ . Therefore

$$\left. \frac{dz}{dt} \right|_{t=0} = (0 + 3)(2 \cos 0) + (0 + 0)(-\sin 0) = 6$$





## ■ The Chain Rule: Case 2

We now consider the situation where  $z = f(x, y)$  but each of  $x$  and  $y$  is a function of two variables  $s$  and  $t$ :  $x = g(s, t)$ ,  $y = h(s, t)$ . Then  $z$  is indirectly a function of  $s$  and  $t$  and we wish to find  $\partial z/\partial s$  and  $\partial z/\partial t$ . Recall that in computing  $\partial z/\partial t$  we hold  $s$  fixed and compute the ordinary derivative of  $z$  with respect to  $t$ . Therefore we can apply Theorem 1 to obtain

$$\frac{\partial z}{\partial t} = \frac{\partial z}{\partial x} \frac{\partial x}{\partial t} + \frac{\partial z}{\partial y} \frac{\partial y}{\partial t}$$

A similar argument holds for  $\partial z/\partial s$  and so we have proved the following version of the Chain Rule.

**2 The Chain Rule (Case 2)** Suppose that  $z = f(x, y)$  is a differentiable function of  $x$  and  $y$ , where  $x = g(s, t)$  and  $y = h(s, t)$  are differentiable functions of  $s$  and  $t$ . Then

$$\frac{\partial z}{\partial s} = \frac{\partial z}{\partial x} \frac{\partial x}{\partial s} + \frac{\partial z}{\partial y} \frac{\partial y}{\partial s} \qquad \frac{\partial z}{\partial t} = \frac{\partial z}{\partial x} \frac{\partial x}{\partial t} + \frac{\partial z}{\partial y} \frac{\partial y}{\partial t}$$

**EXAMPLE 3** If  $z = e^x \sin y$ , where  $x = st^2$  and  $y = s^2t$ , find  $\partial z/\partial s$  and  $\partial z/\partial t$ .

**SOLUTION** Applying Case 2 of the Chain Rule, we get

$$\frac{\partial z}{\partial s} = \frac{\partial z}{\partial x} \frac{\partial x}{\partial s} + \frac{\partial z}{\partial y} \frac{\partial y}{\partial s} = (e^x \sin y)(t^2) + (e^x \cos y)(2st)$$

$$\frac{\partial z}{\partial t} = \frac{\partial z}{\partial x} \frac{\partial x}{\partial t} + \frac{\partial z}{\partial y} \frac{\partial y}{\partial t} = (e^x \sin y)(2st) + (e^x \cos y)(s^2)$$

If we wish, we can now express  $\partial z/\partial s$  and  $\partial z/\partial t$  solely in terms of  $s$  and  $t$  by substituting  $x = st^2$ ,  $y = s^2t$ , to get

$$\frac{\partial z}{\partial s} = t^2 e^{st^2} \sin(s^2t) + 2ste^{st^2} \cos(s^2t)$$

$$\frac{\partial z}{\partial t} = 2ste^{st^2} \sin(s^2t) + s^2 e^{st^2} \cos(s^2t)$$



## ■ The Chain Rule: General Version

Now we consider the general situation in which a dependent variable  $u$  is a function of  $n$  intermediate variables  $x_1, \dots, x_n$ , each of which is, in turn, a function of  $m$  independent variables  $t_1, \dots, t_m$ . Notice that there are  $n$  terms, one for each intermediate variable. The proof is similar to that of Case 1.

**3 The Chain Rule (General Version)** Suppose that  $u$  is a differentiable function of the  $n$  variables  $x_1, x_2, \dots, x_n$  and each  $x_j$  is a differentiable function of the  $m$  variables  $t_1, t_2, \dots, t_m$ . Then  $u$  is a function of  $t_1, t_2, \dots, t_m$  and

$$\frac{\partial u}{\partial t_i} = \frac{\partial u}{\partial x_1} \frac{\partial x_1}{\partial t_i} + \frac{\partial u}{\partial x_2} \frac{\partial x_2}{\partial t_i} + \dots + \frac{\partial u}{\partial x_n} \frac{\partial x_n}{\partial t_i}$$

for each  $i = 1, 2, \dots, m$ .

**EXAMPLE 4** Write out the Chain Rule for the case where  $w = f(x, y, z, t)$  and  $x = x(u, v)$ ,  $y = y(u, v)$ ,  $z = z(u, v)$ , and  $t = t(u, v)$ .

**SOLUTION** We apply Theorem 3 with  $n = 4$  and  $m = 2$ . Figure 3 shows the tree diagram. Although we haven't written the derivatives on the branches, it's understood that if a branch leads from  $y$  to  $u$ , then the partial derivative for that branch is  $\partial y/\partial u$ . With the aid of the tree diagram, we can now write the required expressions:

$$\begin{aligned}\frac{\partial w}{\partial u} &= \frac{\partial w}{\partial x} \frac{\partial x}{\partial u} + \frac{\partial w}{\partial y} \frac{\partial y}{\partial u} + \frac{\partial w}{\partial z} \frac{\partial z}{\partial u} + \frac{\partial w}{\partial t} \frac{\partial t}{\partial u} \\ \frac{\partial w}{\partial v} &= \frac{\partial w}{\partial x} \frac{\partial x}{\partial v} + \frac{\partial w}{\partial y} \frac{\partial y}{\partial v} + \frac{\partial w}{\partial z} \frac{\partial z}{\partial v} + \frac{\partial w}{\partial t} \frac{\partial t}{\partial v}\end{aligned}$$

**EXAMPLE 5** If  $u = x^4y + y^2z^3$ , where  $x = rse^t$ ,  $y = rs^2e^{-t}$ , and  $z = r^2s \sin t$ , find the value of  $\partial u/\partial s$  when  $r = 2$ ,  $s = 1$ ,  $t = 0$ .

**SOLUTION** With the help of the tree diagram in Figure 4, we have

$$\begin{aligned}\frac{\partial u}{\partial s} &= \frac{\partial u}{\partial x} \frac{\partial x}{\partial s} + \frac{\partial u}{\partial y} \frac{\partial y}{\partial s} + \frac{\partial u}{\partial z} \frac{\partial z}{\partial s} \\ &= (4x^3y)(re^t) + (x^4 + 2yz^3)(2rse^{-t}) + (3y^2z^2)(r^2 \sin t)\end{aligned}$$

When  $r = 2$ ,  $s = 1$ , and  $t = 0$ , we have  $x = 2$ ,  $y = 2$ , and  $z = 0$ , so

$$\frac{\partial u}{\partial s} = (64)(2) + (16)(4) + (0)(0) = 192$$

## ■ Implicit Differentiation

The Chain Rule can be used to give a more complete description of the process of implicit differentiation that was introduced in Sections 3.5 and 14.3. We suppose that an equation of the form  $F(x, y) = 0$  defines  $y$  implicitly as a differentiable function of  $x$ , that is,  $y = f(x)$ , where  $F(x, f(x)) = 0$  for all  $x$  in the domain of  $f$ . If  $F$  is differentiable, we can apply Case 1 of the Chain Rule to differentiate both sides of the equation  $F(x, y) = 0$  with respect to  $x$ . Since both  $x$  and  $y$  are functions of  $x$ , we obtain

$$\frac{\partial F}{\partial x} \frac{dx}{dx} + \frac{\partial F}{\partial y} \frac{dy}{dx} = 0$$

But  $dx/dx = 1$ , so if  $\partial F/\partial y \neq 0$  we solve for  $dy/dx$  and obtain

5

$$\frac{dy}{dx} = -\frac{\frac{\partial F}{\partial x}}{\frac{\partial F}{\partial y}} = -\frac{F_x}{F_y}$$

To derive this equation we assumed that  $F(x, y) = 0$  defines  $y$  implicitly as a function of  $x$ . The **Implicit Function Theorem**, proved in advanced calculus, gives conditions under which this assumption is valid: it states that if  $F$  is defined on a disk containing  $(a, b)$ , where  $F(a, b) = 0$ ,  $F_y(a, b) \neq 0$ , and  $F_x$  and  $F_y$  are continuous on the disk, then the equation  $F(x, y) = 0$  defines  $y$  as a function of  $x$  near the point  $(a, b)$  and the derivative of this function is given by Equation 5.

**EXAMPLE 8** Find  $y'$  if  $x^3 + y^3 = 6xy$ .

**SOLUTION** The given equation can be written as

$$F(x, y) = x^3 + y^3 - 6xy = 0$$

so Equation 5 gives

$$\frac{dy}{dx} = -\frac{F_x}{F_y} = -\frac{3x^2 - 6y}{3y^2 - 6x} = -\frac{x^2 - 2y}{y^2 - 2x}$$



Now we suppose that  $z$  is given implicitly as a function  $z = f(x, y)$  by an equation of the form  $F(x, y, z) = 0$ . This means that  $F(x, y, f(x, y)) = 0$  for all  $(x, y)$  in the domain of  $f$ . If  $F$  and  $f$  are differentiable, then we can use the Chain Rule to differentiate the equation  $F(x, y, z) = 0$  as follows:

$$\frac{\partial F}{\partial x} \frac{\partial x}{\partial x} + \frac{\partial F}{\partial y} \frac{\partial y}{\partial x} + \frac{\partial F}{\partial z} \frac{\partial z}{\partial x} = 0$$

But  $\frac{\partial}{\partial x}(x) = 1$  and  $\frac{\partial}{\partial x}(y) = 0$

so this equation becomes

$$\frac{\partial F}{\partial x} + \frac{\partial F}{\partial z} \frac{\partial z}{\partial x} = 0$$

If  $\partial F/\partial z \neq 0$ , we solve for  $\partial z/\partial x$  and obtain the first formula in Equations 6. The formula for  $\partial z/\partial y$  is obtained in a similar manner.

6

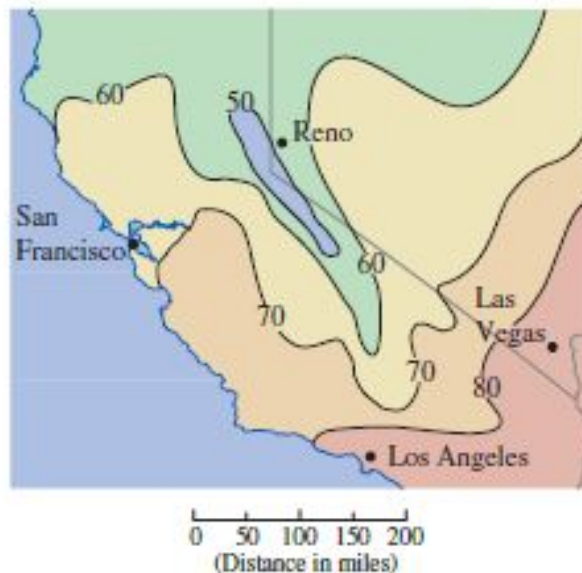
$$\frac{\partial z}{\partial x} = -\frac{\frac{\partial F}{\partial x}}{\frac{\partial F}{\partial z}} = -\frac{F_x}{F_z} \quad \frac{\partial z}{\partial y} = -\frac{\frac{\partial F}{\partial y}}{\frac{\partial F}{\partial z}} = -\frac{F_y}{F_z}$$

Again, a version of the **Implicit Function Theorem** stipulates conditions under which our assumption is valid: if  $F$  is defined within a sphere containing  $(a, b, c)$ , where  $F(a, b, c) = 0$ ,  $F_z(a, b, c) \neq 0$ , and  $F_x$ ,  $F_y$ , and  $F_z$  are continuous inside the sphere, then the equation  $F(x, y, z) = 0$  defines  $z$  as a function of  $x$  and  $y$  near the point  $(a, b, c)$  and this function is differentiable, with partial derivatives given by (6).



## Directional Derivatives and the Gradient Vector

The weather map in Figure 1 shows a contour map of the temperature function  $T(x, y)$  for the states of California and Nevada at 3:00 PM on a day in October. The level curves, or isothermals, join locations with the same temperature. The partial derivative  $T_x$  at a location such as Reno is the rate of change of temperature with respect to distance if we travel east from Reno;  $T_y$  is the rate of change of temperature if we travel north. But what if we want to know the rate of change of temperature when we travel southeast (toward Las Vegas), or in some other direction? In this section we introduce a type of derivative, called a *directional derivative*, that enables us to find the rate of change of a function of two or more variables in any direction.



## ■ Directional Derivatives

Recall that if  $z = f(x, y)$ , then the partial derivatives  $f_x$  and  $f_y$  are defined as

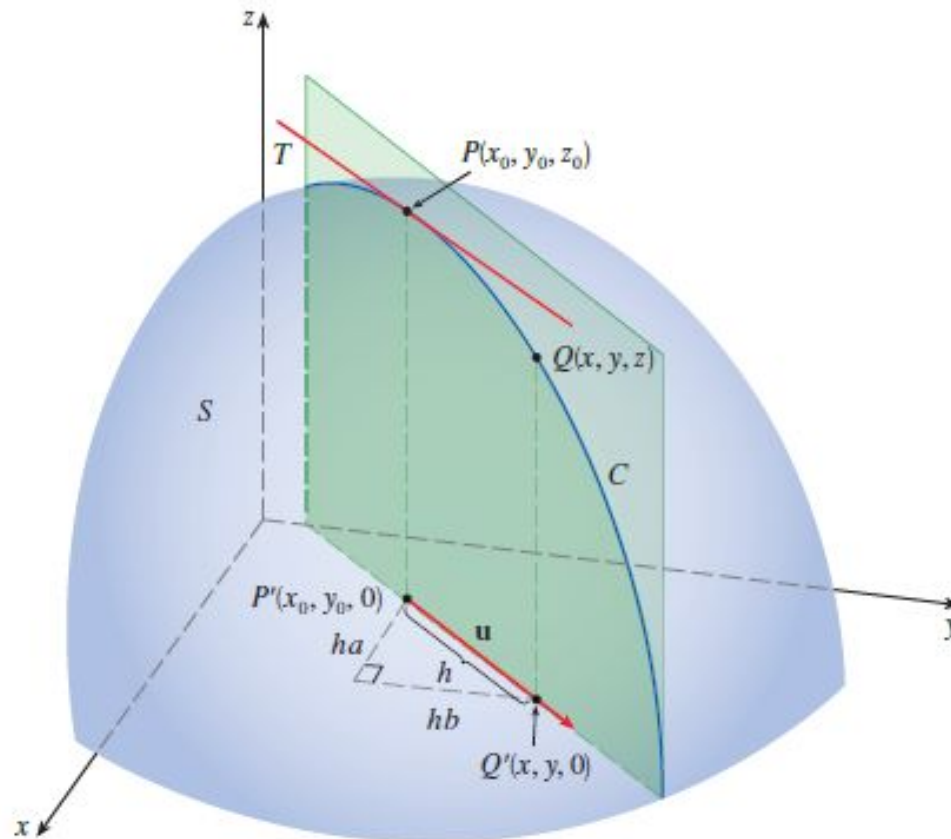
$$f_x(x_0, y_0) = \lim_{h \rightarrow 0} \frac{f(x_0 + h, y_0) - f(x_0, y_0)}{h}$$

1

$$f_y(x_0, y_0) = \lim_{h \rightarrow 0} \frac{f(x_0, y_0 + h) - f(x_0, y_0)}{h}$$

and represent the rates of change of  $z$  in the  $x$ - and  $y$ -directions, that is, in the directions of the unit vectors  $\mathbf{i}$  and  $\mathbf{j}$ .

Suppose that we now wish to find the rate of change of  $z$  at  $(x_0, y_0)$  in the direction of an arbitrary unit vector  $\mathbf{u} = \langle a, b \rangle$ . (See Figure 2.) To do this we consider the surface  $S$  with the equation  $z = f(x, y)$  (the graph of  $f$ ) and we let  $z_0 = f(x_0, y_0)$ . Then the point  $P(x_0, y_0, z_0)$  lies on  $S$ . The vertical plane that passes through  $P$  in the direction of  $\mathbf{u}$  intersects  $S$  in a curve  $C$ . (See Figure 3.) The slope of the tangent line  $T$  to  $C$  at the point  $P$  is the rate of change of  $z$  in the direction of  $\mathbf{u}$ .



If  $Q(x, y, z)$  is another point on  $C$  and  $P', Q'$  are the projections of  $P, Q$  onto the  $xy$ -plane, then the vector  $\overrightarrow{P'Q'}$  is parallel to  $\mathbf{u}$  and so

$$\overrightarrow{P'Q'} = h\mathbf{u} = \langle ha, hb \rangle$$

for some scalar  $h$ . Therefore  $x - x_0 = ha$ ,  $y - y_0 = hb$ , so  $x = x_0 + ha$ ,  $y = y_0 + hb$ , and

$$\frac{\Delta z}{h} = \frac{z - z_0}{h} = \frac{f(x_0 + ha, y_0 + hb) - f(x_0, y_0)}{h}$$

If we take the limit as  $h \rightarrow 0$ , we obtain the rate of change of  $z$  (with respect to distance) in the direction of  $\mathbf{u}$ , which is called the directional derivative of  $f$  in the direction of  $\mathbf{u}$ .

**2 Definition** The directional derivative of  $f$  at  $(x_0, y_0)$  in the direction of a unit vector  $\mathbf{u} = \langle a, b \rangle$  is

$$D_{\mathbf{u}}f(x_0, y_0) = \lim_{h \rightarrow 0} \frac{f(x_0 + ha, y_0 + hb) - f(x_0, y_0)}{h}$$

if this limit exists.

By comparing Definition 2 with Equations 1, we see that if  $\mathbf{u} = \mathbf{i} = \langle 1, 0 \rangle$ , then  $D_{\mathbf{i}}f = f_x$  and if  $\mathbf{u} = \mathbf{j} = \langle 0, 1 \rangle$ , then  $D_{\mathbf{j}}f = f_y$ . In other words, the partial derivatives of  $f$  with respect to  $x$  and  $y$  are just special cases of the directional derivative.

When we compute the directional derivative of a function defined by a formula, we generally use the following theorem.

**3 Theorem** If  $f$  is a differentiable function of  $x$  and  $y$ , then  $f$  has a directional derivative in the direction of any unit vector  $\mathbf{u} = \langle a, b \rangle$  and

$$D_{\mathbf{u}}f(x, y) = f_x(x, y)a + f_y(x, y)b$$

**PROOF** If we define a function  $g$  of the single variable  $h$  by

$$g(h) = f(x_0 + ha, y_0 + hb)$$

then, by the definition of a derivative, we have

$$\begin{aligned} \mathbf{4} \quad g'(0) &= \lim_{h \rightarrow 0} \frac{g(h) - g(0)}{h} = \lim_{h \rightarrow 0} \frac{f(x_0 + ha, y_0 + hb) - f(x_0, y_0)}{h} \\ &= D_{\mathbf{u}}f(x_0, y_0) \end{aligned}$$

On the other hand, we can write  $g(h) = f(x, y)$ , where  $x = x_0 + ha$ ,  $y = y_0 + hb$ , so Case 1 of the Chain Rule (Theorem 14.5.1) gives

$$g'(h) = \frac{\partial f}{\partial x} \frac{dx}{dh} + \frac{\partial f}{\partial y} \frac{dy}{dh} = f_x(x, y) a + f_y(x, y) b$$

If we now put  $h = 0$ , then  $x = x_0$ ,  $y = y_0$ , and

$$\boxed{5} \quad g'(0) = f_x(x_0, y_0) a + f_y(x_0, y_0) b$$

Comparing Equations 4 and 5, we see that

$$D_{\mathbf{u}} f(x_0, y_0) = f_x(x_0, y_0) a + f_y(x_0, y_0) b \quad \blacksquare$$

If the unit vector  $\mathbf{u}$  makes an angle  $\theta$  with the positive  $x$ -axis (as in Figure 5), then we can write  $\mathbf{u} = \langle \cos \theta, \sin \theta \rangle$  and the formula in Theorem 3 becomes

$$\boxed{6} \quad D_{\mathbf{u}} f(x, y) = f_x(x, y) \cos \theta + f_y(x, y) \sin \theta$$

**EXAMPLE 2** Find the directional derivative  $D_{\mathbf{u}}f(x, y)$  if

$$f(x, y) = x^3 - 3xy + 4y^2$$

and  $\mathbf{u}$  is the unit vector in the direction given by angle  $\theta = \pi/6$ , measured from the positive  $x$ -axis. What is  $D_{\mathbf{u}}f(1, 2)$ ?

**SOLUTION** Formula 6 gives

$$\begin{aligned} D_{\mathbf{u}}f(x, y) &= f_x(x, y) \cos \frac{\pi}{6} + f_y(x, y) \sin \frac{\pi}{6} \\ &= (3x^2 - 3y) \frac{\sqrt{3}}{2} + (-3x + 8y) \frac{1}{2} \\ &= \frac{1}{2} [3\sqrt{3}x^2 - 3x + (8 - 3\sqrt{3})y] \end{aligned}$$

Therefore

$$D_{\mathbf{u}}f(1, 2) = \frac{1}{2} [3\sqrt{3}(1)^2 - 3(1) + (8 - 3\sqrt{3})(2)] = \frac{13 - 3\sqrt{3}}{2} \quad \blacksquare$$

## ■ The Gradient Vector

Notice from Theorem 3 that the directional derivative of a differentiable function can be written as the dot product of two vectors:

$$\begin{aligned} \boxed{7} \quad D_{\mathbf{u}}f(x, y) &= f_x(x, y)a + f_y(x, y)b \\ &= \langle f_x(x, y), f_y(x, y) \rangle \cdot \langle a, b \rangle \\ &= \langle f_x(x, y), f_y(x, y) \rangle \cdot \mathbf{u} \end{aligned}$$

The first vector in this dot product occurs not only in computing directional derivatives but in many other contexts as well. So we give it a special name (the *gradient* of  $f$ ) and a special notation ( $\mathbf{grad} f$  or  $\nabla f$ , which is read “del  $f$ ”).

**8 Definition** If  $f$  is a function of two variables  $x$  and  $y$ , then the **gradient** of  $f$  is the vector function  $\nabla f$  defined by

$$\nabla f(x, y) = \langle f_x(x, y), f_y(x, y) \rangle = \frac{\partial f}{\partial x} \mathbf{i} + \frac{\partial f}{\partial y} \mathbf{j}$$



**EXAMPLE 3** If  $f(x, y) = \sin x + e^{xy}$ , then

$$\nabla f(x, y) = \langle f_x, f_y \rangle = \langle \cos x + ye^{xy}, xe^{xy} \rangle$$

and

$$\nabla f(0, 1) = \langle 2, 0 \rangle$$



With this notation for the gradient vector, we can rewrite Equation 7 for the directional derivative of a differentiable function as

9

$$D_{\mathbf{u}}f(x, y) = \nabla f(x, y) \cdot \mathbf{u}$$

This expresses the directional derivative in the direction of a unit vector  $\mathbf{u}$  as the scalar projection of the gradient vector onto  $\mathbf{u}$ .