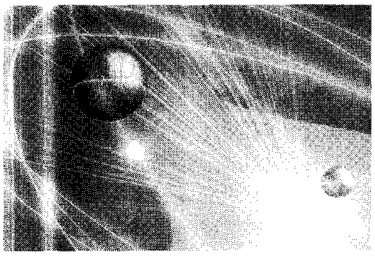

CHAPTER 12

Partial Differentiation

Introduction This chapter is concerned with extending the idea of the derivative to real functions of a vector variable, that is, to functions depending on several real variables.

12.1 Functions of Several Variables

The notation $y = f(x)$ is used to indicate that the variable y depends on the single real variable x, that is, that y is a function of x. The domain of such a function f is a set of real numbers. Many quantities can be regarded as depending on more than one real variable and thus to be functions of more than one variable. For example, the volume of a circular cylinder of radius r and height h is given by $V = \pi r^2 h$; we say that V is a function of the two variables r and h. If we choose to denote this function by f, then we would write $V = f(r, h)$ where

$$f(r,h) = \pi r^2 h, \qquad (r \geq 0, \quad h \geq 0).$$

Thus, f is a function of two variables having as *domain* the set of points in the rh-plane with coordinates (r, h) satisfying $r > 0$ and $h > 0$. Similarly, the relationship $w = f(x, y, z) = x + 2y - 3z$ defines w as a function of the three variables x, y, and z, with domain the whole of $\mathbb{R}^3$, or, if we state explicitly, some particular subset of $\mathbb{R}^3$.

By analogy with the corresponding definition for functions of one variable, we define a function of n variables as follows:

DEFINITION 1

A **function** f of n real variables is a rule that assigns a *unique* real number $f(x_1, x_2, \dots, x_n)$ to each point $(x_1, x_2, \dots, x_n)$ in some subset $\mathcal{D}(f)$ of $\mathbb{R}^n$. $\mathcal{D}(f)$ is called the **domain** of f. The set of real numbers $f(x_1, x_2, \dots, x_n)$ obtained from points in the domain is called the **range** of f.

As for functions of one variable, the **domain convention** specifies that the domain of a function of n variables is the largest set of points $(x_1, x_2, \dots, x_n)$ for which $f(x_1, x_2, \dots, x_n)$ makes sense as a real number, unless that domain is explicitly stated to be a smaller set.

Most of the examples we consider hereafter will be functions of two or three independent variables. When a function f depends on two variables, we will usually call these independent variables x and y, and we will use z to denote the dependent variable that represents the value of the function; that is, $z = f(x, y)$. We will normally use x, y, and z as the independent variables of a function of three variables and w as the value of the function: $w = f(x, y, z)$. Some definitions will be given, and some theorems will be stated (and proved) only for the two-variable case, but extensions to three or more variables will usually be obvious.

Graphical Representations

The graph of a *function* f of one variable (i.e., the graph of the *equation* $y = f(x)$) is the set of points in the xy-plane having coordinates $(x, f(x))$, where x is in the domain of f. Similarly, the graph of a *function* f of two variables (the graph of the *equation* $z = f(x, y)$) is the set of points in 3-space having coordinates $(x, y, f(x, y))$, where (x, y) belongs to the domain of f. This graph is a surface in $\mathbb{R}^3$ lying above (if $f(x, y) > 0$) or below (if $f(x, y) < 0$) the domain of f in the xy-plane. (See Figure 12.1.) The graph of a function of three variables is a three-dimensional *hypersurface* in 4-space, $\mathbb{R}^4$. In general, the graph of a function of n variables is an n-dimensional *surface* in $\mathbb{R}^{n+1}$. We will not attempt to *draw* graphs of functions of more than two variables!

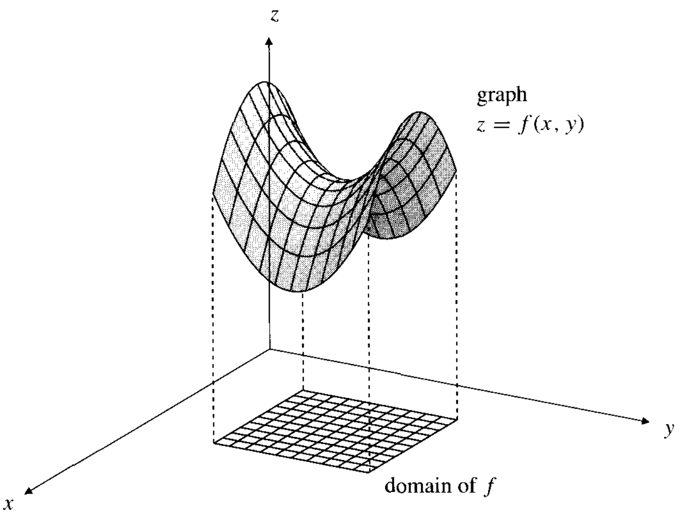

Figure 12.1 The graph of $f(x, y)$ is the surface with equation $z = f(x, y)$ defined for points (x, y) in the domain of f

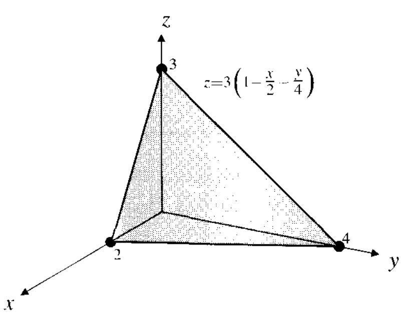

Figure 12.2

Example 1 Consider the function

$$f(x, y) = 3\left(1 - \frac{x}{2} - \frac{y}{4}\right), \qquad (0 \le x \le 2, \quad 0 \le y \le 4 - 2x).$$

The graph of f is the plane triangular surface with vertices at $(2, 0, 0)$, $(0, 4, 0)$, and $(0, 0, 3)$. (See Figure 12.2.) If the domain of f had not been explicitly stated to be a particular set in the xy-plane, the graph would have been the whole plane through these three points.

Example 2 Consider $f(x, y) = \sqrt{9 - x^2 - y^2}$. The expression under the square root cannot be negative, so the domain is the disk $x^2 + y^2 \le 9$ in the xy-plane.

If we square the equation $z = \sqrt{9 - x^2 - y^2}$, we can rewrite the result in the form $x^2 + y^2 + z^2 = 9$. This is a sphere of radius 3 centred at the origin. However, the graph of f is only the upper hemisphere where $z \ge 0$. (See Figure 12.3.)

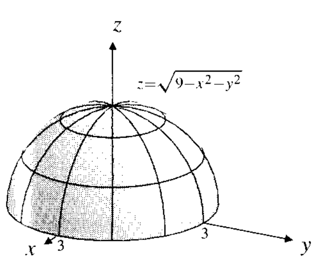

Figure 12.3

Since it is necessary to project the surface $z = f(x, y)$ onto a two-dimensional page, most such graphs are difficult to sketch without considerable artistic talent and training. Nevertheless, you should always try to visualize such a graph and sketch it as best you can. Sometimes it is convenient to sketch only part of a graph, for instance, the part lying in the first octant. It is also helpful to determine (and sketch) the intersections of the graph with various planes, especially the coordinate planes, and planes parallel to the coordinate planes. (See Figure 12.1.)

Some mathematical software packages will produce plots of three-dimensional graphs to help you get a feeling for how the corresponding functions behave. Figure 12.1 is an example of such a computer-drawn graph, as is Figure 12.4 below. Along with most of the other mathematical graphics in this book, both were produced using the mathematical graphics software package **MG**. Later in this section we discuss how to use Maple to produce such graphs.

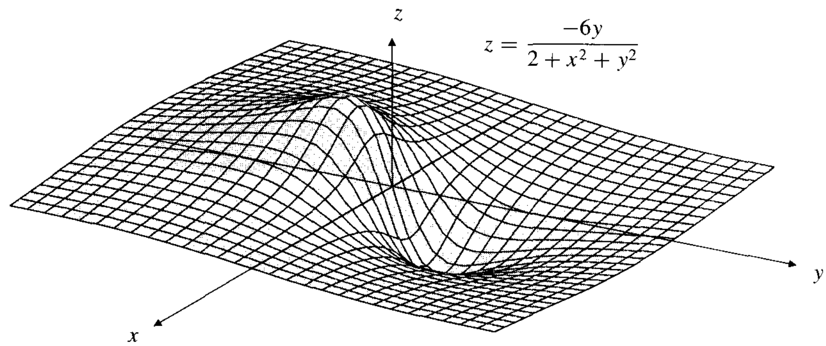

Figure 12.4 The graph of $z = \dfrac{-6y}{2 + x^2 + y^2}$

Another way to represent the function $f(x, y)$ graphically is to produce a two-dimensional *topographic map* of the surface $z = f(x, y)$. In the xy-plane we sketch the curves $f(x, y) = C$ for various values of the constant C. These curves are called **level curves** of f because they are the vertical projections onto the xy-plane of the curves in which the graph $z = f(x, y)$ intersects the horizontal (level) planes $z = C$. The graph and some level curves of the function $f(x, y) = x^2 + y^2$ are shown in Figure 12.5. The graph is a circular paraboloid in 3-space; the level curves are circles centred at the origin in the xy-plane.

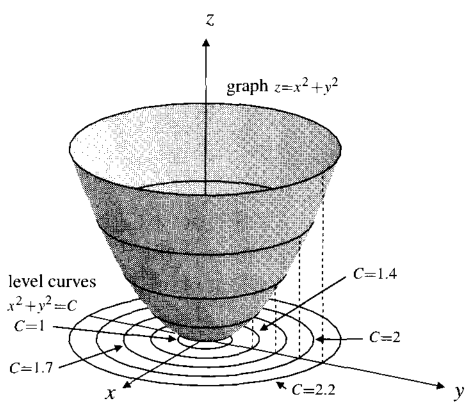

Figure 12.5 The graph of $f(x, y) = x^2 + y^2$ and some level curves of f

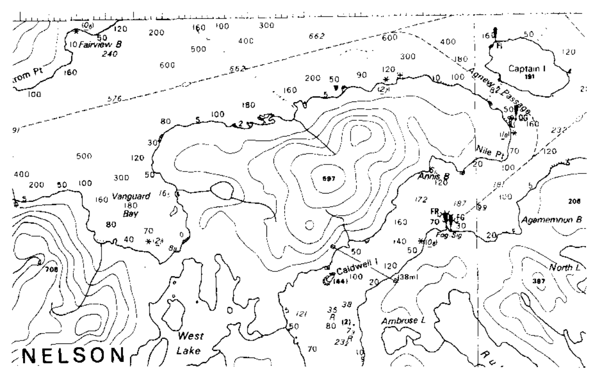

Figure 12.6 Level curves (contours) representing elevation in a topographic map

The contour curves in the topographic map in Figure 12.6 show the elevations, in 100 m increments above sea level, on part of Nelson Island on the British Columbia coast. Since these contours are drawn for equally spaced values of C, the spacing of the contours themselves conveys information about the relative steepness at various places on the mountains; the land is steepest where the contour lines are closest together. Observe also that the streams shown cross the contours at right angles.

They take the route of steepest descent. Isotherms (curves of constant temperature) and isobars (curves of constant pressure) on weather maps are also examples of level curves.

Example 3 The level curves of the function $f(x, y) = 3\left(1 - \frac{x}{2} - \frac{y}{4}\right)$ of Example 1 are the segments of the straight lines

$$3\left(1 - \frac{x}{2} - \frac{y}{4}\right) = C \quad \text{or} \quad \frac{x}{2} + \frac{y}{4} = 1 - \frac{C}{3}, \quad (0 \le C \le 3),$$

which lie in the first quadrant. Several such level curves are shown in Figure 12.7(a). They correspond to equally spaced values of C, and their equal spacing indicates the uniform steepness of the graph of f in Figure 12.2.

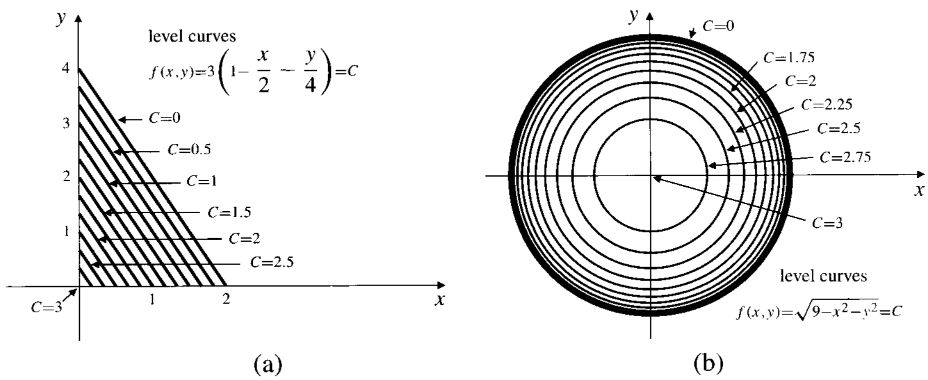

Figure 12.7
(a) Level curves of $3\left(1 - \frac{x}{2} - \frac{y}{4}\right)$
(b) Level curves of $\sqrt{9 - x^2 - y^2}$

Example 4 The level curves of the function $f(x, y) = \sqrt{9 - x^2 - y^2}$ of Example 2 are the concentric circles

$$\sqrt{9 - x^2 - y^2} = C \quad \text{or} \quad x^2 + y^2 = 9 - C^2, \quad (0 \le C \le 3).$$

Observe the spacing of these circles in Figure 12.7(b); they are plotted for several equally spaced values of C. The bunching of the circles as $C \to 0+$ indicates the steepness of the hemispherical surface that is the graph of f. (See Figure 12.3.)

A function determines its level curves with any given spacing between consecutive values of C. However, level curves only determine the function if *all of them* are known.

Example 5 The level curves of the function $f(x, y) = x^2 - y^2$ are the curves $x^2 - y^2 = C$. For $C = 0$ the level "curve" is the pair of straight lines $x = y$ and $x = -y$. For other values of C the level curves are rectangular hyperbolas with these lines as asymptotes. (See Figure 12.8(a).) The graph of f is the saddle-like hyperbolic paraboloid in Figure 12.8(b).

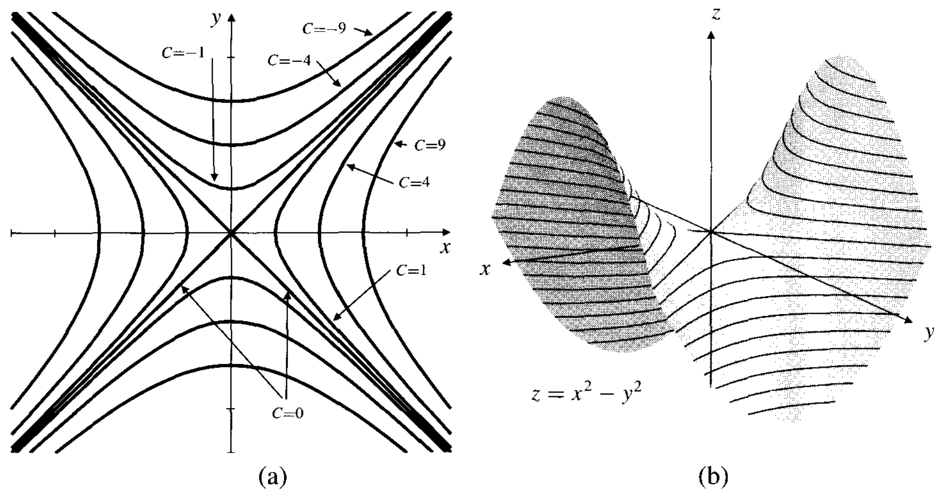

Figure 12.8
(a) Level curves of $x^2 - y^2$
(b) The graph of $x^2 - y^2$

Example 6 Describe and sketch some level curves of the function $z = g(x, y)$ defined by $z \geq 0$, $x^2 + (y - z)^2 = 2z^2$. Also sketch the graph of g.

Solution The level curve $z = g(x, y) = C$ (where C is a positive constant) has equation $x^2 + (y - C)^2 = 2C^2$ and is, therefore, a circle of radius $\sqrt{2}C$ centred at $(0, C)$. Level curves for C in increments of 0.1 from 0 to 1 are shown in Figure 12.9(a). These level curves intersect rays from the origin at equal spacing (the spacing is different for different rays) indicating that the surface $z = g(x, y)$ is an oblique circular cone. See Figure 12.9(b). ∎

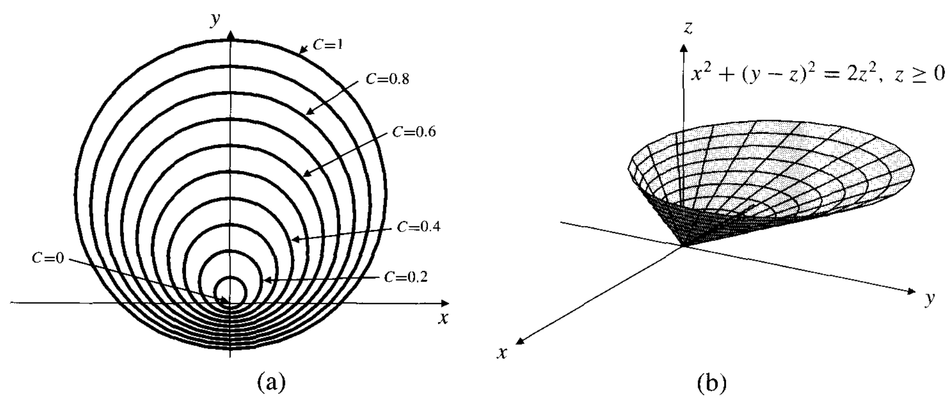

Figure 12.9
(a) Level curves of $z = g(x, y)$ for Example 6
(b) The graph of $z = g(x, y)$

Although the *graph* of a function $f(x, y, z)$ of three variables cannot easily be drawn (it is a three-dimensional *hypersurface* in 4-space), such a function has **level surfaces** in 3-space that can, perhaps, be drawn. These level surfaces have equations $f(x, y, z) = C$ for various choices of the constant C. For instance, the level surfaces of the function $f(x, y, z) = x^2 + y^2 + z^2$ are concentric spheres centred at the origin. Figure 12.10 shows a few level surfaces of the function $f(x, y, z) = x^2 - z$. They are parabolic cylinders.

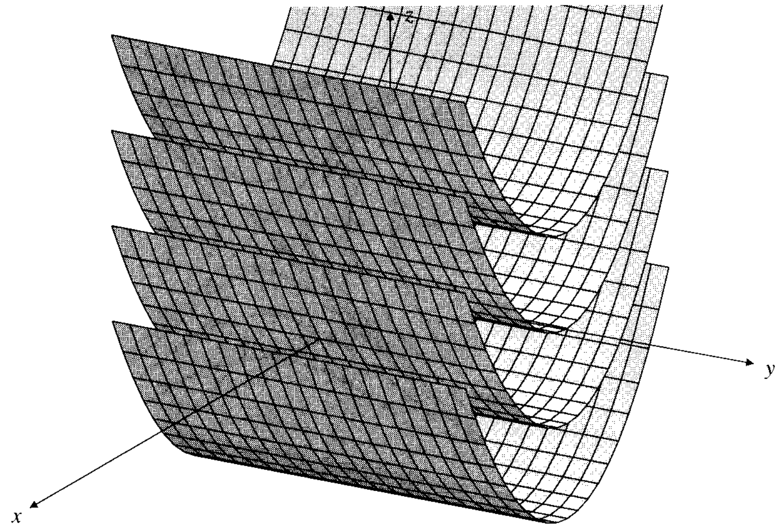

Figure 12.10 Level surfaces of $f(x, y, z) = x^2 - z$

Using Maple Graphics

Like many mathematical software packages, Maple has several plotting routines to help you visualize the behaviour of functions of two and three variables. We mention only a few of them here; there are many more. Most of the plotting routines are in the **plots** package, so you should begin any Maple session where you want to use them with the input

```
> with(plots):
```

To save space, we won't show any of the plot output here. You will need to play with modifications to the various plot commands to obtain the kind of output you desire.

The graph of a function $f(x, y)$ of two variables (or an expression in x and y) can be plotted over a rectangle in the xy-plane with a call to the **plot3d** routine. For example,

```
> f := -6*y/(2+x^2+y^2);
> plot3d(f, x=-6..6, y=-6..6);
```

will plot a surface similar to the one in Figure 12.4 but without axes and viewed from a steeper angle. You can add many kinds of options to the command to change the output. For instance,

```
> plot3d(f, x=-6..6, y=-6..6, axes=boxed,
    orientation=[30,70]);
```

will plot the same surface within a 3-dimensional rectangular box with scales on three of its edges indicating the coordinate values. (If we had said `axes=normal` instead, we would have got the usual coordinate axes through the origin, but they tend to be harder to see against the background of the surface, so `axes=boxed` is usually preferable. The option `orientation=[30,70]` results in the plot's being viewed from the direction making angle 70° with the z-axis and lying in a plane containing the z-axis making an angle 30° with the xz-plane. (The default value of the orientation is [45, 45] if the option is not specified.) By default, the surface plotted by plot3d is ruled by two families of curves, representing its intersection with vertical planes $x = a$ and $y = b$ for several equally spaced values of a and b, and it is coloured opaquely so that hidden parts do not show.

Instead of plot3d, you can use **contourplot3d** to get a plot of the surface ruled by contours on which the value of the function is constant. If you don't get enough contours by default, you can include a `contours=n` option to specify the number you want.

```
> contourplot3d(f, x=-6..6, y=-6..6, axes=boxed,
   contours=24);
```

The contours are the projections of the level curves onto the graph of the surface. Alternatively, you can get a two-dimensional plot of the level curves themselves using **contourplot**

```
> contourplot(f, x=-6..6, y=-6..6, axes=normal,
   contours=24);
```

Other options you may want to include with plot3d or contourplot3d are:

(a) `view=zmin..zmax` to specify the range of values of the function (i.e., z) to show in the plot.

(b) `grid=[m,n]` to specify the number of x and y values at which to evaluate the function. If your plot doesn't look smooth enough, try $m = n = 20$ or 30 or even higher values.

The graph of an equation, $f(x, y) = 0$, in the xy-plane can be generated without solving the equation for x or y first, by using **implicitplot**.

```
> implicitplot(x^3-y^2-5*x*y-x-5, x=-6..7, y=-5..6);
```

will produce the graph of $x^3 - y^2 - 5xy - x - 5 = 0$ on the rectangle $-6 \le x \le 7$, $-5 \le y \le 6$. There is also an **implicitplot3d** routine to plot the surface in 3-space having an equation of the form $f(x, y, z) = 0$. For this routine you must specify ranges for all three variables;

```
> implicitplot3d(x^2+y^2-z^2-1, x=-4..4, y=-4..4,
   z=-3..3, axes=boxed);
```

plots the hyperboloid $z^2 = x^2 + y^2 - 1$.

Finally, we observe that Maple is no more capable than we are of drawing graphs of functions of three or more variables, since it doesn't have four-dimensional plot capability. The best we can do is plot a set of level surfaces for such a function:

```
> implicitplot3d({z-x^2-2,z-x^2,z-x^2+2},x=-2..2,
   y=-2..2, z=-2..5, axes=boxed);
```

It is possible to construct a sequence of *plot structures* and assign them to, say, the elements of a list variable, without actually plotting them. Then all the plots can be plotted simultaneously using the **display** function.

```
> for c from -1 to 1 do
   p[c] := implicitplot3d(z^2-x^2-y^2-2*c, x=-3..3,
   y=-3..3, z=0..2,
   color=COLOR(RGB,(1+c)/2,(1-c)/2,1)) od:
> display([seq(p[c],c=-1..1)], axes=boxed,
   orientation=[30,40]);
```

Note that the command creating the plots is terminated with a colon ":" rather than the usual semicolon. If you don't suppress the output in this way, you will get vast amounts of meaningless numerical output as the plots are constructed. The color=... option is an attempt to give the three plots a different colour so they

Specify the domains of the functions in Exercises 1–10.

1. $f(x, y) = \dfrac{x+y}{x-y}$ **2.** $f(x, y) = \sqrt{xy}$

3. $f(x, y) = \dfrac{x}{x^2+y^2}$ **4.** $f(x, y) = \dfrac{xy}{x^2-y^2}$

5. $f(x, y) = \sqrt{4x^2+9y^2-36}$

6. $f(x, y) = \dfrac{1}{\sqrt{x^2-y^2}}$ **7.** $f(x, y) = \ln(1+xy)$

8. $f(x, y) = \sin^{-1}(x+y)$

9. $f(x, y, z) = \dfrac{xyz}{x^2+y^2+z^2}$

10. $f(x, y, z) = \dfrac{e^{xyz}}{\sqrt{xyz}}$

Sketch the graphs of the functions in Exercises 11–18.

11. $f(x, y) = x, \quad (0 \le x \le 2, \quad 0 \le y \le 3)$

12. $f(x, y) = \sin x, \quad (0 \le x \le 2\pi, \quad 0 \le y \le 1)$

13. $f(x, y) = y^2, \quad (-1 \le x \le 1, \quad -1 \le y \le 1)$

14. $f(x, y) = 4 - x^2 - y^2, \quad (x^2+y^2 \le 4,\ x \ge 0,\ y \ge 0)$

15. $f(x, y) = \sqrt{x^2+y^2}$ **16.** $f(x, y) = 4 - x^2$

17. $f(x, y) = |x| + |y|$ **18.** $f(x, y) = 6 - x - 2y$

Sketch some of the level curves of the functions in Exercises 19–26.

19. $f(x, y) = x - y$ **20.** $f(x, y) = x^2 + 2y^2$

21. $f(x, y) = xy$ **22.** $f(x, y) = \dfrac{x^2}{y}$

23. $f(x, y) = \dfrac{x-y}{x+y}$ **24.** $f(x, y) = \dfrac{y}{x^2+y^2}$

25. $f(x, y) = xe^{-y}$ **26.** $f(x, y) = \sqrt{\dfrac{1}{y} - x^2}$

Exercises 27–28 refer to Figure 12.11, which shows contours of a hilly region with heights given in metres.

27. At which of the points A and B is the landscape steeper? How do you know?

28. Describe the topography of the region near point C.

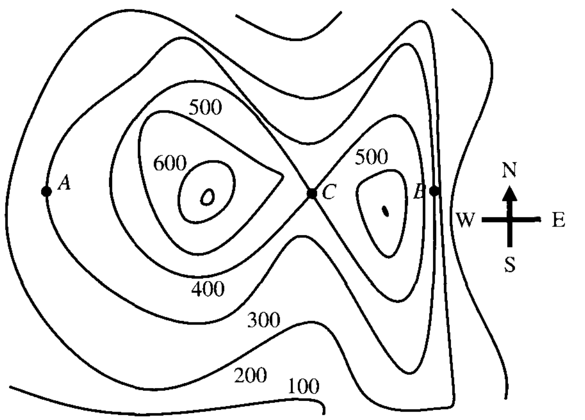

Figure 12.11

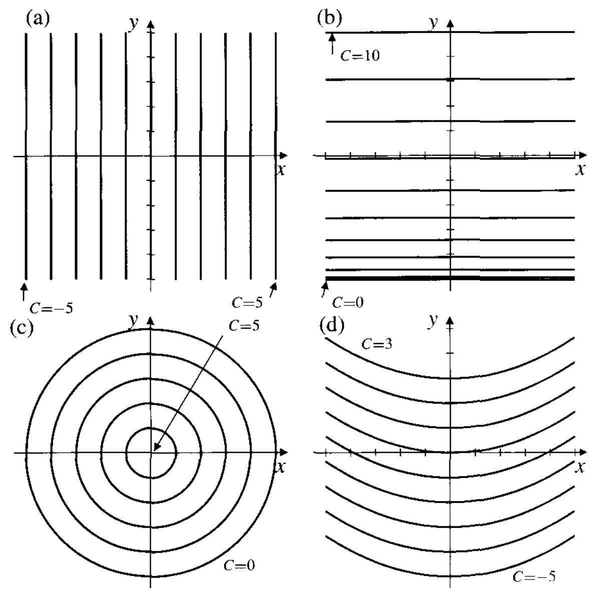

Figure 12.12

Describe the graphs of the functions $f(x, y)$ for which families of level curves $f(x, y) = C$ are shown in the figures referred to in Exercises 29–32. Assume that each family corresponds to equally spaced values of C and that the behaviour of the family is representative of all such families for the function.

29. See Figure 12.12(a). **30.** See Figure 12.12(b).

31. See Figure 12.12(c). **32.** See Figure 12.12(d).

33. Are the curves $y = (x - C)^2$ level curves of a function $f(x, y)$? What property must a family of curves in a region of the xy-plane have to be the family of level curves of a function defined in the region?

34. If we assume $z \geq 0$, the equation $4z^2 = (x - z)^2 + (y - z)^2$ defines z as a function of x and y. Sketch some level curves of this function. Describe its graph.

35. Find $f(x, y)$ if each level curve $f(x, y) = C$ is a circle centred at the origin and having radius
(a) C (b) C^2 (c) $\sqrt{C}$ (d) $\ln C$.

36. Find $f(x, y, z)$ if for each constant C the level surface $f(x, y, z) = C$ is a plane having intercepts C^3, $2C^3$, and $3C^3$ on the x-axis, the y-axis, and the z-axis, respectively.

Describe the level surfaces of the functions specified in Exercises 37–41.

37. $f(x, y, z) = x^2 + y^2 + z^2$

38. $f(x, y, z) = x + 2y + 3z$

39. $f(x, y, z) = x^2 + y^2$

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

41. $f(x, y, z) = |x| + |y| + |z|$

42. Describe the "level hypersurfaces" of the function

$$f(x, y, z, t) = x^2 + y^2 + z^2 + t^2.$$

Use Maple or other computer graphing software to plot the graphs and the level curves of the functions in Exercises 43–48.

43. $\dfrac{1}{1 + x^2 + y^2}$

44. $\dfrac{\cos x}{1 + y^2}$

45. $\dfrac{y}{1 + x^2 + y^2}$

46. $\dfrac{x}{(x^2 - 1)^2 + y^2}$

47. xy

48. $\dfrac{1}{xy}$

12.2 Limits and Continuity

Before reading this section you should review the concepts of neighbourhood, open and closed sets, and boundary and interior points introduced in Section 10.1.

The concept of the limit of a function of several variables is similar to that for functions of one variable. For clarity we present the definition for functions of two variables only; the general case is similar.

We might say that $f(x, y)$ approaches the limit L as the point (x, y) approaches the point (a, b), and write

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

if all points of a neighbourhood of (a, b), except possibly the point (a, b) itself, belong to the domain of f, and if $f(x, y)$ approaches L as (x, y) approaches (a, b). However, it is more convenient to define the limit in such a way that (a, b) can be a boundary point of the domain of f. Thus, our formal definition will generalize the one-dimensional notion of one-sided limit as well.

DEFINITION 2

Definition of Limit

We say that $\lim_{(x,y)\to(a,b)} f(x, y) = L$, provided that

(i) every neighbourhood of (a, b) contains points of the domain of f different from (a, b), and

(ii) for every positive number ϵ there exists a positive number $\delta = \delta(\epsilon)$ such that $|f(x, y) - L| < \epsilon$ holds whenever (x, y) is in the domain of f and satisfies $0 < \sqrt{(x - a)^2 + (y - b)^2} < \delta$.

Condition (i) is included in Definition 2 because it is not appropriate to consider limits at *isolated* points of the domain of f, that is, points with neighbourhoods that contain no other points of the domain.

If a limit exists it is unique. For a single-variable function f, the existence of $\lim_{x \to a} f(x)$ implies that $f(x)$ approaches the same finite number as x approaches a from either the right or the left. Similarly, for a function of two variables, $\lim_{(x,y)\to(a,b)} f(x, y) = L$ exists only if $f(x, y)$ approaches the same number L no matter how (x, y) approaches (a, b) in the domain of f. In particular, (x, y) can approach (a, b) along any curve that lies in $\mathcal{D}(f)$. It is not necessary that $L = f(a, b)$ even if $f(a, b)$ is defined. The examples below illustrate these assertions.

All the usual laws of limits extend to functions of several variables in the obvious way. For example, if $\lim_{(x,y)\to(a,b)} f(x, y) = L$, $\lim_{(x,y)\to(a,b)} g(x, y) = M$, and every neighbourhood of (a, b) contains points in $\mathcal{D}(f) \cap \mathcal{D}(g)$ other than (a, b), then

$$\lim_{(x,y)\to(a,b)} \big(f(x, y) \pm g(x, y)\big) = L \pm M,$$

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

$$\lim_{(x,y)\to(a,b)} \frac{f(x, y)}{g(x, y)} = \frac{L}{M}, \qquad \text{provided } M \neq 0.$$

Also, if $F(t)$ is continuous at $t = L$, then

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

Example 1 (a) $\displaystyle\lim_{(x,y)\to(2,3)} 2x - y^2 = 4 - 9 = -5,$

(b) $\displaystyle\lim_{(x,y)\to(a,b)} x^2 y = a^2 b,$

(c) $\displaystyle\lim_{(x,y)\to(\pi/3,2)} y \sin\left(\frac{x}{y}\right) = 2 \sin\left(\frac{\pi}{6}\right) = 1.$ ■

Example 2 The function $f(x, y) = \sqrt{1 - x^2 - y^2}$ is continuous at all points of its domain, the *closed* disk $x^2 + y^2 \leq 1$. Of course, (x, y) can approach points of the bounding circle $x^2 + y^2 = 1$ only from within the disk. ■

The following examples show that the requirement that $f(x, y)$ approach the same limit *no matter how* (x, y) approaches (a, b) can be very restrictive, and makes limits in two or more variables much more subtle than in the single-variable case.

Example 3 Investigate the limiting behaviour of $f(x, y) = \dfrac{2xy}{x^2 + y^2}$ as (x, y) approaches $(0, 0)$.

Solution Note that $f(x, y)$ is defined at all points of the xy-plane except the origin $(0, 0)$. We can still ask whether $\lim_{(x,y)\to(0,0)} f(x, y)$ exists or not. If we let (x, y) approach $(0, 0)$ along the x-axis $(y = 0)$, then $f(x, y) = f(x, 0) \to 0$ (because $f(x, 0) = 0$ identically). Thus, $\lim_{(x,y)\to(0,0)} f(x, y)$ must be 0 if it exists at all. Similarly, at all points of the y-axis we have $f(x, y) = f(0, y) = 0$.

However, at points of the line $x = y$, f has a different constant value; $f(x, x) = 1$. Since the limit of $f(x, y)$ is 1 as (x, y) approaches $(0, 0)$ along this line, it follows that $f(x, y)$ cannot have a unique limit at the origin. That is,

$$\lim_{(x,y)\to(0,0)} \frac{2xy}{x^2 + y^2} \qquad \text{does not exist.}$$

Observe that $f(x, y)$ has a constant value on any ray from the origin (on the ray $y = kx$ the value is $2k/(1 + k^2)$), but these values differ on different rays. The level curves of f are the rays from the origin (with the origin itself removed). It is difficult to sketch the graph of f near the origin. The first-octant part of the graph is the "hood-shaped" surface in Figure 12.13(a). ■

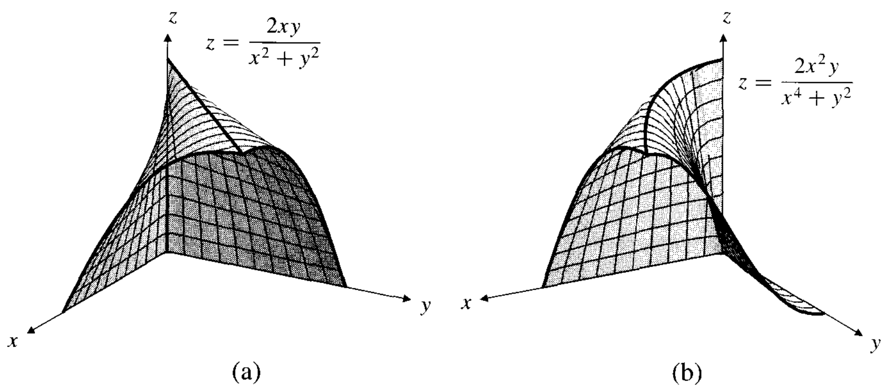

Figure 12.13

(a) $f(x, y)$ has different limits as $(x, y) \to (0, 0)$ along different straight lines

(b) $f(x, y)$ has the same limit 0 as $(x, y) \to (0, 0)$ along any straight line but has limit 1 as $(x, y) \to (0, 0)$ along $y = x^2$

Example 4 Investigate the limiting behaviour of $f(x, y) = \dfrac{2x^2y}{x^4 + y^2}$ as (x, y) approaches $(0, 0)$.

Solution As in Example 3, $f(x, y)$ vanishes identically on the coordinate axes, so $\lim_{(x,y)\to(0,0)} f(x, y)$ must be 0 if it exists at all. If we examine $f(x, y)$ at points of the ray $y = kx$, we obtain

$$f(x, kx) = \frac{2kx^3}{x^4 + k^2x^2} = \frac{2kx}{x^2 + k^2} \to 0, \qquad \text{as} \quad x \to 0 \quad (k \neq 0).$$

Thus, $f(x, y) \to 0$ as $(x, y) \to (0, 0)$ along *any* straight line through the origin. We might be tempted to conclude, therefore, that $\lim_{(x,y)\to(0,0)} f(x, y) = 0$, but this is incorrect. Observe the behaviour of $f(x, y)$ along the curve $y = x^2$:

$$f(x, x^2) = \frac{2x^4}{x^4 + x^4} = 1.$$

Thus, $f(x, y)$ does not approach 0 as (x, y) approaches the origin along this curve, so $\lim_{(x,y)\to(0,0)} f(x, y)$ does not exist. The level curves of f are pairs of parabolas of the form $y = kx^2$, $y = x^2/k$ with the origin removed. See Figure 12.13(b) for the first octant part of the graph of f. ■

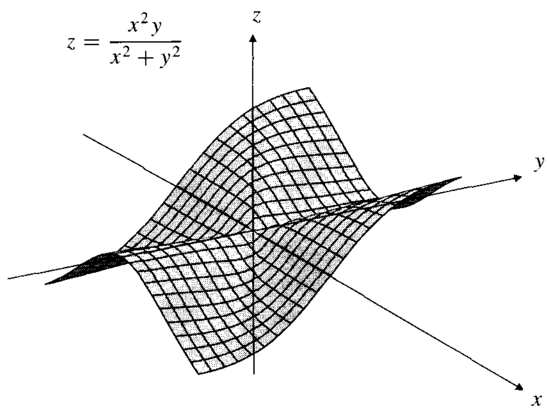

Figure 12.14
$$\lim_{(x,y)\to(0,0)} \frac{x^2y}{x^2+y^2} = 0$$

Example 5 Show that the function $f(x, y) = \dfrac{x^2y}{x^2+y^2}$ does have a limit at the origin; specifically,

$$\lim_{(x,y)\to(0,0)} \frac{x^2y}{x^2+y^2} = 0.$$

Solution This function is also defined everywhere except at the origin. Observe that since $x^2 \le x^2 + y^2$, we have

$$|f(x, y) - 0| = \left|\frac{x^2y}{x^2+y^2}\right| \le |y| \le \sqrt{x^2+y^2},$$

which approaches zero as $(x, y) \to (0, 0)$. (See Figure 12.14.) Formally, if $\epsilon > 0$ is given and we take $\delta = \epsilon$, then $|f(x, y) - 0| < \epsilon$ whenever $0 < \sqrt{x^2+y^2} < \delta$, so $f(x, y)$ has limit 0 as $(x, y) \to (0, 0)$ by Definition 2. ■

As for functions of one variable, continuity of a function f at a point of its domain is defined directly in terms of the limit.

DEFINITION 3

The function $f(x, y)$ is continuous at the point (a, b) if

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

It remains true that sums, differences, products, quotients, and compositions of continuous functions are continuous. The functions of Examples 3 and 4 above are continuous wherever they are defined, that is, at all points except the origin. There is no way to define $f(0, 0)$ so that these functions become continuous at the origin. They show that the continuity of the single-variable functions $f(x, b)$ at $x = a$ and $f(a, y)$ at $y = b$ does *not* imply that $f(x, y)$ is continuous at (a, b). In fact, even if $f(x, y)$ is continuous along every straight line through (a, b), it still need not be continuous at (a, b). (See Exercises 16–17 below.) Note, however, that the function $f(x, y)$ of Example 5, although not defined at the origin, has a continuous extension to that point. If we extend the domain of f by defining

$f(0,0) = \lim_{(x,y)\to(0,0)} f(x,y) = 0$, then f is continuous on the whole xy-plane.

As for functions of one variable, the existence of a limit of a function at a point does not imply that the function is continuous at that point. The function

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

satisfies $\lim_{(x,y)\to(0,0)} f(x,y) = 0$, which is not equal to $f(0,0)$, so f is not continuous at $(0,0)$. Of course, we can *make* f continuous at $(0,0)$ by redefining its value at that point to be 0.

Exercises 12.2

In Exercises 1–12, evaluate the indicated limit or explain why it does not exist.

1. $\lim_{(x,y)\to(2,-1)} xy + x^2$

2. $\lim_{(x,y)\to(0,0)} \sqrt{x^2+y^2}$

3. $\lim_{(x,y)\to(0,0)} \dfrac{x^2+y^2}{y}$

4. $\lim_{(x,y)\to(0,0)} \dfrac{x}{x^2+y^2}$

5. $\lim_{(x,y)\to(1,\pi)} \dfrac{\cos(xy)}{1-x-\cos y}$

6. $\lim_{(x,y)\to(0,1)} \dfrac{x^2(y-1)^2}{x^2+(y-1)^2}$

7. $\lim_{(x,y)\to(0,0)} \dfrac{y^3}{x^2+y^2}$

8. $\lim_{(x,y)\to(0,0)} \dfrac{\sin(x-y)}{\cos(x+y)}$

9. $\lim_{(x,y)\to(0,0)} \dfrac{\sin(xy)}{x^2+y^2}$

10. $\lim_{(x,y)\to(1,2)} \dfrac{2x^2-xy}{4x^2-y^2}$

11. $\lim_{(x,y)\to(0,0)} \dfrac{x^2y^2}{x^2+y^4}$

12. $\lim_{(x,y)\to(0,0)} \dfrac{x^2y^2}{2x^4+y^4}$

13. How can the function

$$f(x,y) = \frac{x^2+y^2-x^3y^3}{x^2+y^2}, \qquad (x,y) \neq (0,0),$$

be defined at the origin so that it becomes continuous at all points of the xy-plane?

14. How can the function

$$f(x,y) = \frac{x^3-y^3}{x-y}, \qquad (x \neq y),$$

be defined along the line $x = y$ so that the resulting function is continuous on the whole xy-plane?

15. What is the domain of

$$f(x,y) = \frac{x-y}{x^2-y^2}?$$

Does $f(x,y)$ have a limit as $(x,y) \to (1,1)$? Can the domain of f be extended so that the resulting function is continuous at $(1,1)$? Can the domain be extended so that the resulting function is continuous everywhere in the xy-plane?

*** 16.** Given a function $f(x,y)$ and a point (a,b) in its domain, define single-variable functions g and h as follows:

$$g(x) = f(x,b), \qquad h(y) = f(a,y).$$

If g is continuous at $x = a$ and h is continuous at $y = b$, does it follow that f is continuous at (a,b)? Conversely, does the continuity of f at (a,b) guarantee the continuity of g at a and the continuity of h at b? Justify your answers.

*** 17.** Let $\mathbf{u} = u\mathbf{i} + v\mathbf{j}$ be a unit vector, and let

$$f_{\mathbf{u}}(t) = f(a+tu, b+tv)$$

be the single-variable function obtained by restricting the domain of $f(x,y)$ to points of the straight line through (a,b) parallel to $\mathbf{u}$. If $f_{\mathbf{u}}(t)$ is continuous at $t = 0$ for every unit vector $\mathbf{u}$, does it follow that f is continuous at (a,b)? Conversely, does the continuity of f at (a,b) guarantee the continuity of $f_{\mathbf{u}}(t)$ at $t = 0$? Justify your answers.

*** 18.** What condition must the nonnegative integers m, n, and p satisfy to guarantee that $\lim_{(x,y)\to(0,0)} x^m y^n/(x^2+y^2)^p$ exists? Prove your answer.

*** 19.** What condition must the constants a, b, and c satisfy to guarantee that $\lim_{(x,y)\to(0,0)} xy/(ax^2+bxy+cy^2)$ exists? Prove your answer.

*** 20.** Can the function $f(x,y) = \dfrac{\sin x \sin^3 y}{1-\cos(x^2+y^2)}$ be defined at $(0,0)$ in such a way that it becomes continuous there? If so, how?

21. Use 2- and 3-dimensional mathematical graphing software to examine the graph and level curves of the function $f(x,y)$ of Example 3 on the region $-1 \leq x \leq 1$, $-1 \leq y \leq 1$, $(x,y) \neq (0,0)$. How would you describe the behaviour of the graph near $(x,y) = (0,0)$?

22. Use 2- and 3-dimensional mathematical graphing software to examine the graph and level curves of the function $f(x, y)$ of Example 4 on the region $-1 \le x \le 1$, $-1 \le y \le 1$, $(x, y) \ne (0, 0)$. How would you describe the behaviour of the graph near $(x, y) = (0, 0)$?

23. The graph of a single-variable function $f(x)$ that is continuous on an interval is a curve that has no *breaks* in it there and that intersects any vertical line through a point in the interval exactly once. What analogous statement can you make about the graph of a bivariate function $f(x, y)$ that is continuous on a region of the xy-plane?

12.3 Partial Derivatives

In this section we begin the process of extending the concepts and techniques of single-variable calculus to functions of more than one variable. It is convenient to begin by considering the rate of change of such functions with respect to one variable at a time. Thus, a function of n variables has *n first-order partial derivatives,* one with respect to each of its independent variables. For a function of two variables, we make this precise in the following definition:

DEFINITION 4

The **first partial derivatives** of the function $f(x, y)$ **with respect to the variables** x and y are the functions $f_1(x, y)$ and $f_2(x, y)$ given by

$$f_1(x, y) = \lim_{h \to 0} \frac{f(x+h, y) - f(x, y)}{h},$$

$$f_2(x, y) = \lim_{k \to 0} \frac{f(x, y+k) - f(x, y)}{k},$$

provided these limits exist.

Each of the two partial derivatives is the limit of a Newton quotient in one of the variables. Observe that $f_1(x, y)$ is just the ordinary first derivative of $f(x, y)$ considered as a function of x only, regarding y as a constant parameter. Similarly, $f_2(x, y)$ is the first derivative of $f(x, y)$ considered as a function of y alone, with x held fixed.

Example 1 If $f(x, y) = x^2 \sin y$, then

$$f_1(x, y) = 2x \sin y \qquad \text{and} \qquad f_2(x, y) = x^2 \cos y.$$

■

The subscripts "1" and "2" in the notations for the partial derivatives refer to the "first" and "second" variables of f. For functions of one variable we use the notation f' for the derivative; the *prime* (′) denotes differentiation with respect to the only variable on which f depends. For functions f of two variables we use f_1 or f_2 to show the variable of differentiation. Do not confuse these subscripts with subscripts used for other purposes, for example, to denote the components of vectors.

The partial derivative $f_1(a, b)$ measures the rate of change of $f(x, y)$ with respect to x at $x = a$ while y is held fixed at b. In graphical terms, the surface $z = f(x, y)$ intersects the vertical plane $y = b$ in a curve. If we take horizontal and vertical lines through the point $(0, b, 0)$ as coordinate axes in the plane $y = b$, then the curve has equation $z = f(x, b)$, and its slope at $x = a$ is $f_1(a, b)$. (See Figure 12.15.) Similarly, $f_2(a, b)$ represents the rate of change of f with respect to y at $y = b$ with x held fixed at a. The surface $z = f(x, y)$ intersects the vertical plane $x = a$ in a curve $z = f(a, y)$ whose slope at $y = b$ is $f_2(a, b)$. (See Figure 12.16.)

Various notations can be used to denote the partial derivatives of $z = f(x, y)$ considered as functions of x and y:

Notations for first partial derivatives

$$\frac{\partial z}{\partial x} = \frac{\partial}{\partial x} f(x, y) = f_1(x, y) = D_1 f(x, y)$$
$$\frac{\partial z}{\partial y} = \frac{\partial}{\partial y} f(x, y) = f_2(x, y) = D_2 f(x, y)$$

The symbol "$\partial/\partial x$" should be read as "partial with respect to x" so "$\partial z/\partial x$" is "partial z with respect to x." The reason for distinguishing "∂" from the "d" of ordinary derivatives of single-variable functions will be made clear later.

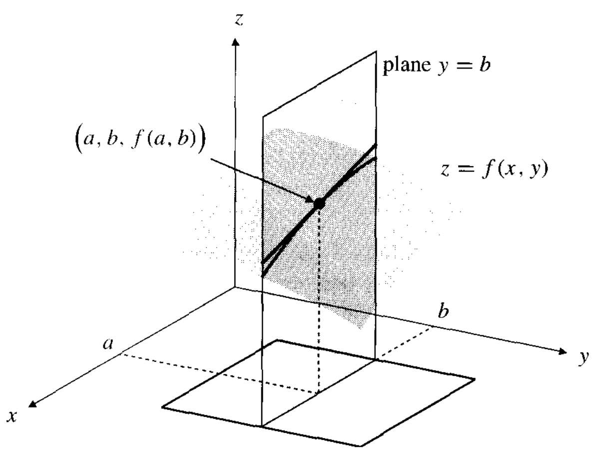

Figure 12.15 $f_1(a, b)$ is the slope of the curve of intersection of $z = f(x, y)$ and the vertical plane $y = b$ at $x = a$

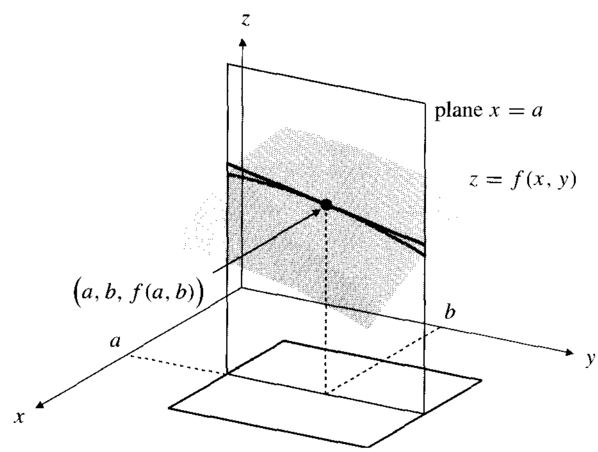

Figure 12.16 $f_2(a, b)$ is the slope of the curve of intersection of $z = f(x, y)$ and the vertical plane $x = a$ at $y = b$

Values of partial derivatives at a particular point (a, b) are denoted similarly:

Values of partial derivatives

$$\left.\frac{\partial z}{\partial x}\right|_{(a,b)} = \left.\left(\frac{\partial}{\partial x} f(x, y)\right)\right|_{(a,b)} = f_1(a, b) = D_1 f(a, b)$$
$$\left.\frac{\partial z}{\partial y}\right|_{(a,b)} = \left.\left(\frac{\partial}{\partial y} f(x, y)\right)\right|_{(a,b)} = f_2(a, b) = D_2 f(a, b)$$

Some authors prefer to use f_x or $\partial f/\partial x$ and f_y or $\partial f/\partial y$ instead of f_1 and f_2. However, this can lead to problems of ambiguity when compositions of functions arise. For instance, suppose $f(x, y) = x^2 y$. By $f_1(x^2, xy)$ we clearly mean

$$\left(\frac{\partial}{\partial x} f(x, y)\right)\bigg|_{(x^2,xy)} = 2xy\bigg|_{(x^2,xy)} = (2)(x^2)(xy) = 2x^3 y.$$

But does $f_x(x^2, xy)$ mean the same thing? One could argue that $f_x(x^2, xy)$ should mean

$$\frac{\partial}{\partial x}\left(f(x^2, xy)\right) = \frac{\partial}{\partial x}\left((x^2)^2(xy)\right) = \frac{\partial}{\partial x}(x^5 y) = 5x^4 y.$$

In order to avoid such ambiguities we usually prefer to use f_1 and f_2 instead of f_x and f_y. (However, in some situations where no confusion is likely to occur we may still use the notations f_x and f_y, and also $\partial f/\partial x$ and $\partial f/\partial y$.)

All the standard differentiation rules for sums, products, reciprocals, and quotients continue to apply to partial derivatives.

Example 2 Find $\partial z/\partial x$ and $\partial z/\partial y$ if $z = x^3 y^2 + x^4 y + y^4$.

Solution $\partial z/\partial x = 3x^2 y^2 + 4x^3 y$ and $\partial z/\partial y = 2x^3 y + x^4 + 4y^3$. ■

Example 3 Find $f_1(0, \pi)$ if $f(x, y) = e^{xy}\cos(x + y)$.

Solution

$$f_1(x, y) = y\, e^{xy}\cos(x + y) - e^{xy}\sin(x + y),$$
$$f_1(0, \pi) = \pi\, e^0\cos(\pi) - e^0\sin(\pi) = -\pi.$$

■

The single-variable version of the Chain Rule also continues to apply to, say, $f\big(g(x, y)\big)$, where f is a function of only one variable having derivative f':

$$\frac{\partial}{\partial x} f\big(g(x, y)\big) = f'\big(g(x, y)\big)\, g_1(x, y), \quad \frac{\partial}{\partial y} f\big(g(x, y)\big) = f'\big(g(x, y)\big)\, g_2(x, y).$$

We will develop versions of the Chain Rule for more complicated compositions of multivariate functions in Section 12.5.

Example 4 If f is an everywhere differentiable function of one variable, show that $z = f(x/y)$ satisfies the *partial differential equation*

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

Solution By the (single-variable) Chain Rule,

$$\frac{\partial z}{\partial x} = f'\left(\frac{x}{y}\right)\left(\frac{1}{y}\right) \quad \text{and} \quad \frac{\partial z}{\partial y} = f'\left(\frac{x}{y}\right)\left(\frac{-x}{y^2}\right).$$

Hence,

$$x\frac{\partial z}{\partial x} + y\frac{\partial z}{\partial y} = f'\left(\frac{x}{y}\right)\left(x \times \frac{1}{y} + y \times \frac{-x}{y^2}\right) = 0.$$

■

Definition 4 can be extended in the obvious way to cover functions of more than two variables. If f is a function of n variables $x_1, x_2, \ldots, x_n$, then f has n first partial derivatives, $f_1(x_1, x_2, \ldots, x_n)$, $f_2(x_1, x_2, \ldots, x_n)$, $\ldots$, $f_n(x_1, x_2, \ldots, x_n)$, one with respect to each variable.

Example 5

$$\frac{\partial}{\partial z}\left(\frac{2xy}{1+xz+yz}\right) = -\frac{2xy}{(1+xz+yz)^2}(x+y).$$

Again, all the standard differentiation rules are applied to calculate partial derivatives.

Remark If a single-variable function $f(x)$ has a derivative $f'(a)$ at $x = a$, then f is necessarily continuous at $x = a$. This property does *not* extend to partial derivatives. Even if all the first partial derivatives of a function of several variables exist at a point, the function may still fail to be continuous at that point. See Exercise 36 below.

Tangent Planes and Normal Lines

If the graph $z = f(x, y)$ is a "smooth" surface near the point P with coordinates $\big(a, b, f(a, b)\big)$, then that graph will have a **tangent plane** and a **normal line** at P. The normal line is the line through P that is perpendicular to the surface; for instance, a line joining a point on a sphere to the centre of the sphere is normal to the sphere. Any nonzero vector that is parallel to the normal line at P is called a normal vector to the surface at P. The tangent plane to the surface $z = f(x, y)$ at P is the plane through P that is perpendicular to the normal line at P.

Let us assume that the surface $z = f(x, y)$ has a *nonvertical* tangent plane (and therefore a *nonhorizontal* normal line) at point P. (Later in this chapter we will state precise conditions that guarantee that the graph of a function has a nonvertical tangent plane at a point.) The tangent plane intersects the vertical plane $y = b$ in a straight line that is tangent at P to the curve of intersection of the surface $z = f(x, y)$ and the plane $y = b$. (See Figures 12.15 and 12.17.) This line has slope $f_1(a, b)$, so it is parallel to the vector $\mathbf{T}_1 = \mathbf{i} + f_1(a, b)\mathbf{k}$. Similarly, the tangent plane intersects the vertical plane $x = a$ in a straight line having slope $f_2(a, b)$. This line is therefore parallel to the vector $\mathbf{T}_2 = \mathbf{j} + f_2(a, b)\mathbf{k}$. It follows that the tangent plane, and therefore the surface $z = f(x, y)$ itself, has normal vector

$$\mathbf{n} = \mathbf{T}_2 \times \mathbf{T}_1 = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 0 & 1 & f_2(a,b) \\ 1 & 0 & f_1(a,b) \end{vmatrix} = f_1(a,b)\mathbf{i} + f_2(a,b)\mathbf{j} - \mathbf{k}.$$

A normal vector to $z = f(x, y)$ at $\big(a, b, f(a, b)\big)$ is

$$\mathbf{n} = f_1(a,b)\mathbf{i} + f_2(a,b)\mathbf{j} - \mathbf{k}.$$

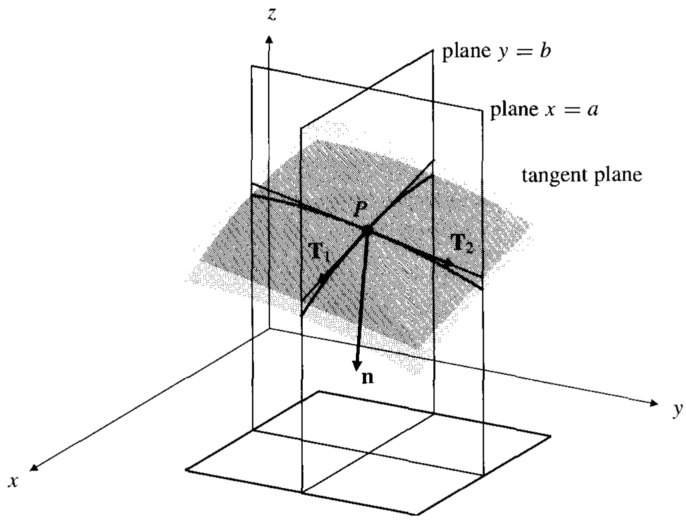

Figure 12.17 The tangent plane and a normal vector to $z = f(x, y)$ at $P = \big(a, b, f(a, b)\big)$

Since the tangent plane passes through $P = (a, b, f(a, b))$, it has equation

$$f_1(a, b)(x - a) + f_2(a, b)(y - b) - (z - f(a, b)) = 0,$$

or, equivalently,

An equation of the tangent plane to $z = f(x, y)$ at $\big(a, b, f(a, b)\big)$ is

$$z = f(a, b) + f_1(a, b)(x - a) + f_2(a, b)(y - b).$$

We shall obtain this result by a different method in Section 12.7.

The normal line to $z = f(x, y)$ at $\big(a, b, f(a, b)\big)$ has direction vector $f_1(a, b)\mathbf{i} + f_2(a, b)\mathbf{j} - \mathbf{k}$ and so has equations

$$\frac{x - a}{f_1(a, b)} = \frac{y - b}{f_2(a, b)} = \frac{z - f(a, b)}{-1}$$

with suitable modifications if either $f_1(a, b) = 0$ or $f_2(a, b) = 0$.

Example 6 Find a normal vector and equations of the tangent plane and normal line to the graph $z = \sin(xy)$ at the point where $x = \pi/3$ and $y = -1$.

Solution The point on the graph has coordinates $(\pi/3, -1, -\sqrt{3}/2)$. Now

$$\frac{\partial z}{\partial x} = y\cos(xy) \qquad \text{and} \qquad \frac{\partial z}{\partial y} = x\cos(xy).$$

At $(\pi/3, -1)$ we have $\partial z/\partial x = -1/2$ and $\partial z/\partial y = \pi/6$. Therefore, the surface has normal vector $\mathbf{n} = -(1/2)\mathbf{i} + (\pi/6)\mathbf{j} - \mathbf{k}$ and tangent plane

$$z = \frac{-\sqrt{3}}{2} - \frac{1}{2}\left(x - \frac{\pi}{3}\right) + \frac{\pi}{6}(y + 1),$$

or, more simply, $3x - \pi y + 6z = 2\pi - 3\sqrt{3}$. The normal line has equation

$$\frac{x-\frac{\pi}{3}}{\frac{-1}{2}} = \frac{y+1}{\frac{\pi}{6}} = \frac{z+\frac{\sqrt{3}}{2}}{-1} \quad \text{or} \quad \frac{6x-2\pi}{-3} = \frac{6y+6}{\pi} = \frac{6z+3\sqrt{3}}{-6}.$$

Example 7 What horizontal plane is tangent to the surface

$$z = x^2 - 4xy - 2y^2 + 12x - 12y - 1,$$

and what is the point of tangency?

Solution A plane is horizontal only if its equation is of the form $z = k$, that is, it is independent of x and y. Therefore, we must have $\partial z/\partial x = \partial z/\partial y = 0$ at the point of tangency. The equations

$$\frac{\partial z}{\partial x} = 2x - 4y + 12 = 0$$
$$\frac{\partial z}{\partial y} = -4x - 4y - 12 = 0$$

have solution $x = -4$, $y = 1$. For these values we have $z = -31$, so the required tangent plane has equation $z = -31$ and the point of tangency is $(-4, 1, -31)$.

Distance from a Point to a Surface: A Geometric Example

Example 8 Find the distance from the point $(3, 0, 0)$ to the hyperbolic paraboloid with equation $z = x^2 - y^2$.

Solution This is an optimization problem of a sort we will deal with in a more systematic way in the next chapter. However, such problems involving minimizing distances from points to surfaces can frequently be solved using geometric methods.

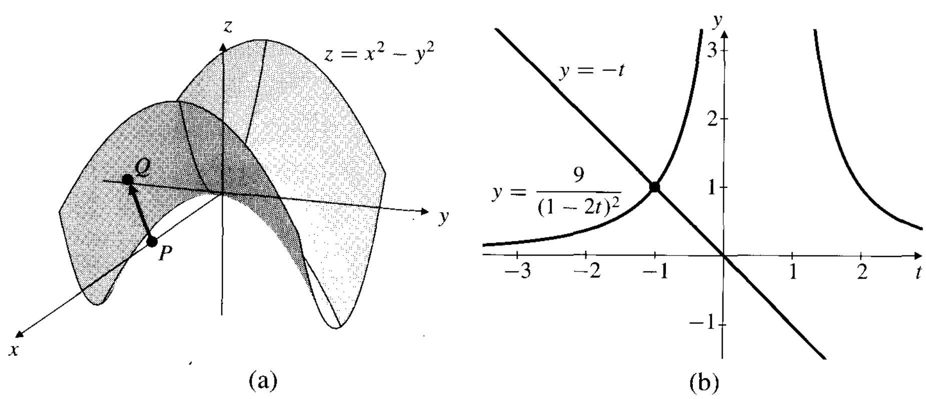

Figure 12.18

(a) If Q is the point on $z = x^2 - y^2$ closest to P, then $\overrightarrow{PQ}$ is normal to the surface

(b) Equation $-t = \dfrac{9}{(1-2t)^2}$ has only one real root, $t = -1$

If $Q = (X, Y, Z)$ is the point on the surface $z = x^2 - y^2$ that is closest to $P = (3, 0, 0)$, then the vector $\overrightarrow{PQ} = (X - 3)\mathbf{i} + Y\mathbf{j} + Z\mathbf{k}$ must be normal to the surface at Q. (See Figure 12.18(a).) Using the partial derivatives of $z = x^2 - y^2$, we know that the vector $\mathbf{n} = 2X\mathbf{i} - 2Y\mathbf{j} - \mathbf{k}$ is normal to the surface at Q. Thus, $\overrightarrow{PQ}$ must be parallel to $\mathbf{n}$, and $\overrightarrow{PQ} = t\mathbf{n}$ for some scalar t. Separated into components, this vector equation states that

$$X - 3 = 2Xt, \qquad Y = -2Yt, \qquad \text{and} \qquad Z = -t.$$

The middle equation implies that either $Y = 0$ or $t = -\frac{1}{2}$. We must consider both of these possibilities.

CASE I If $Y = 0$, then

$$X = \frac{3}{1 - 2t} \qquad \text{and} \qquad Z = -t.$$

But $Z = X^2 - Y^2$, so we must have

$$-t = \frac{9}{(1 - 2t)^2}.$$

This is a cubic equation in t, so we might expect to have to solve it numerically, for instance, by using Newton's Method. However, if we try small integer values of t, we will quickly discover that $t = -1$ is a solution. The graphs of both sides of the equation are shown in Figure 12.18(b). They show that $t = -1$ is the only real solution. Calculating the corresponding values of X and Z, we obtain $(1, 0, 1)$ as a candidate for Q. The distance from this point to P is $\sqrt{5}$.

CASE II If $t = -1/2$, then $X = 3/2$, $Z = 1/2$, and $Y = \pm\sqrt{X^2 - Z} = \pm\sqrt{7}/2$, and the distance from these points to P is $\sqrt{17}/2$.

Since $\frac{17}{4} < 5$, the points $(3/2, \pm\sqrt{7}/2, 1/2)$ are the points on $z = x^2 - y^2$ closest to $(3, 0, 0)$, and the distance from $(3, 0, 0)$ to the surface is $\sqrt{17}/2$ units. ■

Exercises 12.3

In Exercises 1–10, find all the first partial derivatives of the function specified and evaluate them at the given point.

1. $f(x, y) = x - y + 2, \quad (3, 2)$
2. $f(x, y) = xy + x^2, \quad (2, 0)$
3. $f(x, y, z) = x^3y^4z^5, \quad (0, -1, -1)$
4. $g(x, y, z) = \dfrac{xz}{y + z}, \quad (1, 1, 1)$
5. $z = \tan^{-1}\left(\dfrac{y}{x}\right), \quad (-1, 1)$
6. $w = \ln(1 + e^{xyz}), \quad (2, 0, -1)$
7. $f(x, y) = \sin(x\sqrt{y}), \quad \left(\dfrac{\pi}{3}, 4\right)$
8. $f(x, y) = \dfrac{1}{\sqrt{x^2 + y^2}}, \quad (-3, 4)$
9. $w = x^{(y \ln z)}, \quad (e, 2, e)$
10. $g(x_1, x_2, x_3, x_4) = \dfrac{x_1 - x_2^2}{x_3 + x_4^2}, \quad (3, 1, -1, -2)$

In Exercises 11–12, calculate the first partial derivatives of the given functions at $(0, 0)$. You will have to use Definition 4.

11. $f(x, y) = \begin{cases} \dfrac{2x^3 - y^3}{x^2 + 3y^2}, & \text{if } (x, y) \neq (0, 0) \\ 0, & \text{if } (x, y) = (0, 0). \end{cases}$
12. $f(x, y) = \begin{cases} \dfrac{x^2 - 2y^2}{x - y}, & \text{if } x \neq y \\ 0, & \text{if } x = y. \end{cases}$

In Exercises 13–22, find equations of the tangent plane and normal line to the graph of the given function at the point with specified values of x and y.

13. $f(x, y) = x^2 - y^2$ at $(-2, 1)$

14. $f(x, y) = \dfrac{x - y}{x + y}$ at $(1, 1)$

15. $f(x, y) = \cos(x/y)$ at $(\pi, 4)$

16. $f(x, y) = e^{xy}$ at $(2, 0)$

17. $f(x, y) = \dfrac{x}{x^2 + y^2}$ at $(1, 2)$

18. $f(x, y) = y\, e^{-x^2}$ at $(0, 1)$

19. $f(x, y) = \ln(x^2 + y^2)$ at $(1, -2)$

20. $f(x, y) = \dfrac{2xy}{x^2 + y^2}$ at $(0, 2)$

21. $f(x, y) = \tan^{-1}(y/x)$ at $(1, -1)$

22. $f(x, y) = \sqrt{1 + x^3 y^2}$ at $(2, 1)$

23. Find the coordinates of all points on the surface with equation $z = x^4 - 4xy^3 + 6y^2 - 2$ where the surface has a horizontal tangent plane.

24. Find all horizontal planes that are tangent to the surface with equation $z = xye^{-(x^2+y^2)/2}$. At what points are they tangent?

In Exercises 25–31, show that the given function satisfies the given partial differential equation.

◆ **25.** $z = x\, e^y, \quad x \dfrac{\partial z}{\partial x} = \dfrac{\partial z}{\partial y}$

◆ **26.** $z = \dfrac{x + y}{x - y}, \quad x \dfrac{\partial z}{\partial x} + y \dfrac{\partial z}{\partial y} = 0$

◆ **27.** $z = \sqrt{x^2 + y^2}, \quad x \dfrac{\partial z}{\partial x} + y \dfrac{\partial z}{\partial y} = z$

◆ **28.** $w = x^2 + yz, \quad x \dfrac{\partial w}{\partial x} + y \dfrac{\partial w}{\partial y} + z \dfrac{\partial w}{\partial z} = 2w$

◆ **29.** $w = \dfrac{1}{x^2 + y^2 + z^2}, \quad x \dfrac{\partial w}{\partial x} + y \dfrac{\partial w}{\partial y} + z \dfrac{\partial w}{\partial z} = -2w$

◆ **30.** $z = f(x^2 + y^2)$, where f is any differentiable function of one variable,

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

◆ **31.** $z = f(x^2 - y^2)$, where f is any differentiable function of one variable,

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

32. Give a formal definition of the three first partial derivatives of the function $f(x, y, z)$.

33. What is an equation of the "tangent hyperplane" to the graph $w = f(x, y, z)$ at $\big(a, b, c, f(a, b, c)\big)$?

∗ **34.** Find the distance from the point $(1, 1, 0)$ to the circular paraboloid with equation $z = x^2 + y^2$.

∗ **35.** Find the distance from the point $(0, 0, 1)$ to the elliptic paraboloid having equation $z = x^2 + 2y^2$.

∗ **36.** Let $f(x, y) = \begin{cases} \dfrac{2xy}{x^2 + y^2}, & \text{if } (x, y) \neq (0, 0) \\ 0, & \text{if } (x, y) = (0, 0). \end{cases}$

Note that f is not continuous at (0,0). (See Example 3 of Section 12.2.) Therefore its graph is not smooth there. Show, however, that $f_1(0, 0)$ and $f_2(0, 0)$ both exist. Hence, the existence of partial derivatives does not imply that a function of several variables is continuous. This is in contrast to the single-variable case.

37. Determine $f_1(0, 0)$ and $f_2(0, 0)$ if they exist, where

$$f(x, y) = \begin{cases} (x^3 + y) \sin \dfrac{1}{x^2 + y^2}, & \text{if } (x, y) \neq (0, 0) \\ 0, & \text{if } (x, y) = (0, 0). \end{cases}$$

38. Calculate $f_1(x, y)$ for the function in Exercise 37. Is $f_1(x, y)$ continuous at $(0, 0)$?

∗ **39.** Let $f(x, y) = \begin{cases} \dfrac{x^3 - y^3}{x^2 + y^2}, & \text{if } (x, y) \neq (0, 0) \\ 0, & \text{if } (x, y) = (0, 0). \end{cases}$

Calculate $f_1(x, y)$ and $f_2(x, y)$ at all points (x, y) in the plane. Is f continuous at $(0, 0)$? Are f_1 and f_2 continuous at $(0, 0)$?

∗ **40.** Let $f(x, y, z) = \begin{cases} \dfrac{xy^2 z}{x^4 + y^4 + z^4}, & \text{if } (x, y, z) \neq (0, 0, 0) \\ 0, & \text{if } (x, y, z) = (0, 0, 0). \end{cases}$

Find $f_1(0, 0, 0)$, $f_2(0, 0, 0)$, and $f_3(0, 0, 0)$. Is f continuous at $(0, 0, 0)$? Are f_1, f_2, and f_3 continuous at $(0, 0, 0)$?

12.4 Higher-Order Derivatives

Partial derivatives of second and higher orders are calculated by taking partial

$$\frac{\partial^2 z}{\partial x^2} = \frac{\partial}{\partial x}\frac{\partial z}{\partial x} = f_{11}(x, y) = f_{xx}(x, y),$$
$$\frac{\partial^2 z}{\partial y^2} = \frac{\partial}{\partial y}\frac{\partial z}{\partial y} = f_{22}(x, y) = f_{yy}(x, y),$$

and two **mixed** second partial derivatives with respect to x and y,

$$\frac{\partial^2 z}{\partial x \partial y} = \frac{\partial}{\partial x}\frac{\partial z}{\partial y} = f_{21}(x, y) = f_{yx}(x, y),$$
$$\frac{\partial^2 z}{\partial y \partial x} = \frac{\partial}{\partial y}\frac{\partial z}{\partial x} = f_{12}(x, y) = f_{xy}(x, y).$$

Again, we remark that the notations f_{11}, f_{12}, f_{21}, and f_{22} are usually preferable to f_{xx}, f_{xy}, f_{yx}, and f_{yy}, although the latter are often used in partial differential equations. Note that f_{12} indicates differentiation of f *first* with respect to its first variable and *then* with respect to its second variable; f_{21} indicates the opposite order of differentiation. The subscript closest to f indicates which differentiation occurs first.

Similarly, if $w = f(x, y, z)$, then

$$\frac{\partial^5 w}{\partial y \partial x \partial y^2 \partial z} = \frac{\partial}{\partial y}\frac{\partial}{\partial x}\frac{\partial}{\partial y}\frac{\partial}{\partial y}\frac{\partial w}{\partial z} = f_{32212}(x, y, z) = f_{zyyxy}(x, y, z).$$

Example 1 Find the four second partial derivatives of $f(x, y) = x^3y^4$.

Solution

$$f_1(x, y) = 3x^2y^4, \qquad f_2(x, y) = 4x^3y^3,$$
$$f_{11}(x, y) = \frac{\partial}{\partial x}(3x^2y^4) = 6xy^4, \qquad f_{21}(x, y) = \frac{\partial}{\partial x}(4x^3y^3) = 12x^2y^3,$$
$$f_{12}(x, y) = \frac{\partial}{\partial y}(3x^2y^4) = 12x^2y^3, \qquad f_{22}(x, y) = \frac{\partial}{\partial y}(4x^3y^3) = 12x^3y^2.$$

■

Example 2 Calculate $f_{223}(x, y, z)$, $f_{232}(x, y, z)$, and $f_{322}(x, y, z)$ for the function $f(x, y, z) = e^{x-2y+3z}$.

Solution

$$\begin{aligned} f_{223}(x,y,z) &= \frac{\partial}{\partial z}\frac{\partial}{\partial y}\frac{\partial}{\partial y}e^{x-2y+3z} \\ &= \frac{\partial}{\partial z}\frac{\partial}{\partial y}\left(-2e^{x-2y+3z}\right) \\ &= \frac{\partial}{\partial z}\left(4e^{x-2y+3z}\right) = 12\,e^{x-2y+3z}. \end{aligned}$$

$$\begin{aligned} f_{232}(x,y,z) &= \frac{\partial}{\partial y}\frac{\partial}{\partial z}\frac{\partial}{\partial y}e^{x-2y+3z} \\ &= \frac{\partial}{\partial y}\frac{\partial}{\partial z}\left(-2e^{x-2y+3z}\right) \\ &= \frac{\partial}{\partial y}\left(-6e^{x-2y+3z}\right) = 12\,e^{x-2y+3z}. \end{aligned}$$

$$\begin{aligned} f_{322}(x,y,z) &= \frac{\partial}{\partial y}\frac{\partial}{\partial y}\frac{\partial}{\partial z}e^{x-2y+3z} \\ &= \frac{\partial}{\partial y}\frac{\partial}{\partial y}\left(3e^{x-2y+3z}\right) \\ &= \frac{\partial}{\partial y}\left(-6e^{x-2y+3z}\right) = 12\,e^{x-2y+3z}. \end{aligned}$$

■

In both of the examples above observe that the mixed partial derivatives taken with respect to the same variables but in different orders turned out to be equal. This is not a coincidence. It will always occur for sufficiently smooth functions. In particular, the mixed partial derivatives involved are required to be *continuous*. The following theorem presents a more precise statement of this important phenomenon.

THEOREM 1

Equality of mixed partials

Suppose that two mixed nth-order partial derivatives of a function f involve the same differentiations but in different orders. If those partials are continuous at a point P, and if f and all partials of f of order less than n are continuous in a neighbourhood of P, then the two mixed partials are equal at the point P.

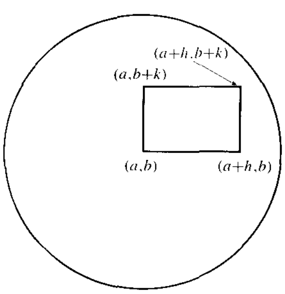

Figure 12.19 A rectangle contained in the disk where f and certain partials are continuous

PROOF We shall prove only a representative special case, showing the equality of $f_{12}(a,b)$ and $f_{21}(a,b)$ for a function f of two variables, provided f_{12} and f_{21} are defined and f_1, f_2, and f are continuous throughout a disk of positive radius centred at (a,b), and f_{12} and f_{21} are continuous at (a,b). Let h and k have sufficiently small absolute values that the point $(a+h,b+k)$ lies in this disk. Then so do all points of the rectangle with sides parallel to the coordinate axes and diagonally opposite corners at (a,b) and $(a+h,b+k)$. (See Figure 12.19.)

Let $Q = f(a+h,b+k) - f(a+h,b) - f(a,b+k) + f(a,b)$ and define single-variable functions $u(x)$ and $v(y)$ by

$$u(x) = f(x,b+k) - f(x,b) \qquad \text{and} \qquad v(y) = f(a+h,y) - f(a,y).$$

Evidently $Q = u(a+h) - u(a)$ and also $Q = v(b+k) - v(b)$. By the (single-variable) Mean-Value Theorem, there exists a number θ_1 satisfying $0 < \theta_1 < 1$ (so that $a + \theta_1 h$ lies between a and $a + h$) such that

$$Q = u(a+h) - u(a) = h\,u'(a+\theta_1 h) = h\left[f_1(a+\theta_1 h, b+k) - f_1(a+\theta_1 h, b)\right].$$

Now we apply the Mean-Value Theorem again, this time to f_1 considered as a function of its second variable, and obtain another number θ_2 satisfying $0 < \theta_2 < 1$ such that

$$f_1(a + \theta_1 h, b + k) - f_1(a + \theta_1 h, b) = k\, f_{12}(a + \theta_1 h, b + \theta_2 k).$$

BEWARE

The Mean-Value Theorem is used four times in this proof, each time to write a difference of the form $g(p + m) - g(p)$ in the form $g'(c)m$, where c is some number between p and $p + m$. It is convenient to write c in the form $p + \theta m$, where θ is some number between 0 and 1.

Thus $Q = hk\, f_{12}(a + \theta_1 h, b + \theta_2 k)$. Two similar applications of the Mean-Value Theorem to $Q = v(b + k) - v(b)$ lead to $Q = hk\, f_{21}(a + \theta_3 h, b + \theta_4 k)$, where θ_3 and θ_4 are two numbers each between 0 and 1. Equating these two expressions for Q and cancelling the common factor hk, we obtain

$$f_{12}(a + \theta_1 h, b + \theta_2 k) = f_{21}(a + \theta_3 h, b + \theta_4 k).$$

Since f_{12} and f_{21} are continuous at (a, b), we can let h and k approach zero to obtain $f_{12}(a, b) = f_{21}(a, b)$, as required.

Exercise 16 below develops an example of a function for which f_{12} and f_{21} exist but are not continuous at $(0, 0)$, and for which $f_{12}(0, 0) \neq f_{21}(0, 0)$.

Remark **Partial Derivatives in Maple** When you use the Maple function **diff** to calculate a derivative, you must include the name of the variable of differentiation. For example, `diff(x^2+y^3, x)` gives the result $2x$. It doesn't matter that the function being differentiated depends on more than one variable since you are telling Maple to differentiate with respect to x. If you wanted the derivative with respect to y, you would input `diff(x^2+y^3,y)` and the output would be $3y^2$. In this context, there is no distinction between ordinary and partial derivatives. There is, however, a difference when you want to apply a *differential operator* to a function f. If f is a function of one variable, you can denote its derivative f' in Maple by `D(f)`. For example,

```
> f := x -> sin(2*x); fprime := D(f);
```

$$f := x \to \sin(2x)$$

$$fprime := x \to 2\cos(2x)$$

The input `fprime(Pi/6)` will now give the output 1, as expected.

If f is a function of two (or more) variables, then `D(f)` no longer makes sense; do we mean f_1 or f_2? We distinguish the two (or more) first partials by using subscripts with D.

```
> f := (x,y) -> exp(3*y)*sin(2*x);
```

$$f := (x, y) \to e^{(3y)} * \sin(2x)$$

```
> fone := D[1](f); ftwo := D[2](f);
```

$$fone := (x, y) \to 2e^{(3y)} * \cos(2x)$$

$$ftwo := (x, y) \to 3e^{(3y)} * \sin(2x)$$

Higher-order partials are denoted with multiple subscripts (within one set of square brackets).

```
> D[1,1,2](f)(Pi/4, 0);
```

$$-12$$

You don't need to worry about the order of the subscripts in a mixed partial. Maple assumes the partials are continuous, even if it doesn't know what the function is. Even if `g` has not been assigned any meaning during the current Maple session, the input `D[1,2](g)(x,y)-D[2,1](g)(x,y);` produces the output 0.

The Laplace and Wave Equations

Many important and interesting phenomena are modelled by functions of several variables that satisfy certain *partial differential equations*. In the following examples we encounter two particular partial differential equations that arise frequently in mathematics and the physical sciences. Exercises 17–19 below introduce another such equation with important applications.

Example 3 Show that for any real number k the functions

$$z = e^{kx}\cos(ky) \qquad \text{and} \qquad z = e^{kx}\sin(ky)$$

satisfy the partial differential equation

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

at every point in the xy-plane.

Solution For $z = e^{kx}\cos(ky)$ we have

$$\frac{\partial z}{\partial x} = k\,e^{kx}\cos(ky), \qquad \frac{\partial z}{\partial y} = -k\,e^{kx}\sin(ky),$$
$$\frac{\partial^2 z}{\partial x^2} = k^2\,e^{kx}\cos(ky), \qquad \frac{\partial^2 z}{\partial y^2} = -k^2\,e^{kx}\cos(ky).$$

Thus,

$$\frac{\partial^2 z}{\partial x^2} + \frac{\partial^2 z}{\partial y^2} = k^2\,e^{kx}\cos(ky) - k^2\,e^{kx}\cos(ky) = 0.$$

The calculation for $z = e^{kx}\sin(ky)$ is similar. ■

Remark The partial differential equation in the above example is called the (two-dimensional) **Laplace equation**. A function of two variables having continuous second partial derivatives in a region of the plane is said to be **harmonic** there if it satisfies Laplace's equation. Such functions play a critical role in the theory of differentiable functions of a *complex variable* and are used to model various physical quantities such as steady-state temperature distributions, fluid flows, and electric and magnetic potential fields. Harmonic functions have many interesting properties. They have derivatives of all orders, and they are *analytic,* that is, they are the sums of their (multivariable) Taylor series. Moreover, a harmonic function can achieve maximum and minimum values only on the boundary of its domain.

Laplace's equation, and therefore harmonic functions, can be considered in any number of dimensions. (See Exercises 13 and 14 below.)

Example 4 If f and g are any twice-differentiable functions of one variable,

$$\frac{\partial^2 w}{\partial t^2} = c^2 \frac{\partial^2 w}{\partial x^2}.$$

Solution Using the Chain Rule for functions of one variable we obtain

$$\frac{\partial w}{\partial t} = -c\, f'(x-ct) + c\, g'(x+ct), \qquad \frac{\partial w}{\partial x} = f'(x-ct) + g'(x+ct),$$

$$\frac{\partial^2 w}{\partial t^2} = c^2\, f''(x-ct) + c^2\, g''(x+ct), \qquad \frac{\partial^2 w}{\partial x^2} = f''(x-ct) + g''(x+ct).$$

Thus w satisfies the given differential equation. ■

Remark The partial differential equation in the above example is called the (one-dimensional) **wave equation**. If t measures time, then $f(x - ct)$ represents a waveform travelling to the right along the x-axis with speed c. (See Figure 12.20.) Similarly, $g(x + ct)$ represents a waveform travelling to the left with speed c. Unlike the solutions of Laplace's equation that must be infinitely differentiable, solutions of the wave equation need only have enough derivatives to satisfy the differential equation. The functions f and g are otherwise arbitrary.

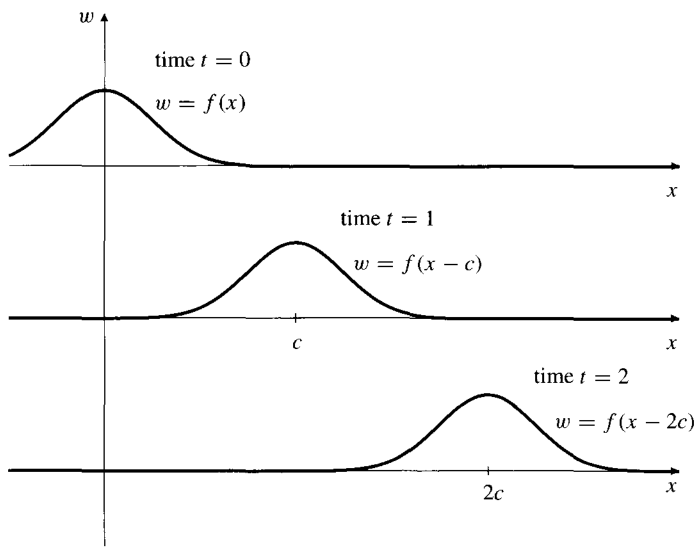

Figure 12.20 $w = f(x - ct)$ represents a waveform moving to the right with speed c

Exercises 12.4

In Exercises 1–6, find all the second partial derivatives of the given function.

1. $z = x^2(1 + y^2)$ **2.** $f(x, y) = x^2 + y^2$

3. $w = x^3y^3z^3$ **4.** $z = \sqrt{3x^2 + y^2}$

5. $z = x\,e^y - y\,e^x$

6. $f(x, y) = \ln\big(1 + \sin(xy)\big)$

* **7.** How many mixed partial derivatives of order 3 can a function of three variables have? If they are all continuous, how many different values can they have at one point? Find the mixed partials of order 3 for $f(x, y, z) = x\,e^{xy}\cos(xz)$ that involve two differentiations with respect to z and one with respect to x.

Show that the functions in Exercises 8–12 are harmonic in the plane regions indicated.

8. $f(x, y) = A(x^2 - y^2) + Bxy$ in the whole plane (A and B are constants.)

9. $f(x, y) = 3x^2y - y^3$ in the whole plane (Can you think of another polynomial of degree 3 in x and y that is also harmonic?)

10. $f(x, y) = \dfrac{x}{x^2 + y^2}$ everywhere except at the origin

11. $f(x, y) = \ln(x^2 + y^2)$ everywhere except at the origin

12. $\tan^{-1}(y/x)$ except at points on the y-axis

13. Show that $w = e^{3x+4y}\sin(5z)$ is harmonic in all of $\mathbb{R}^3$, that is, it satisfies everywhere the 3-dimensional Laplace equation

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

14. Assume that $f(x, y)$ is harmonic in the xy-plane. Show that each of the functions $z\,f(x, y)$, $x\,f(y, z)$, and $y\,f(z, x)$ is harmonic in the whole of $\mathbb{R}^3$. What condition should the constants a, b, and c satisfy to ensure that $f(ax + by, cz)$ is harmonic in $\mathbb{R}^3$?

15. Suppose the functions $u(x, y)$ and $v(x, y)$ have continuous second partial derivatives and satisfy the **Cauchy–Riemann equations**

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \quad \text{and} \quad \frac{\partial v}{\partial x} = -\frac{\partial u}{\partial y}.$$

Show that u and v are both harmonic.

* **16.** Let $F(x, y) = \begin{cases} \dfrac{2xy(x^2 - y^2)}{x^2 + y^2}, & \text{if } (x, y) \neq (0, 0) \\ 0, & \text{if } (x, y) = (0, 0) \end{cases}$

Calculate $F_1(x, y)$, $F_2(x, y)$, $F_{12}(x, y)$, and $F_{21}(x, y)$ at points $(x, y) \neq (0, 0)$. Also calculate these derivatives at $(0, 0)$. Observe that $F_{21}(0, 0) = 2$ and $F_{12}(0, 0) = -2$. Does this result contradict Theorem 1? Explain why.

The heat (diffusion) equation

◆ **17.** Show that the function $u(x, t) = t^{-1/2}\,e^{-x^2/4t}$ satisfies the partial differential equation

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}.$$

This equation is called the **one-dimensional heat equation** because it models heat diffusion in an insulated rod (with $u(x, t)$ representing the temperature at position x at time t) and other similar phenomena.

◆ **18.** Show that the function $u(x, y, t) = t^{-1}\,e^{-(x^2+y^2)/4t}$ satisfies the **two-dimensional heat equation**

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}.$$

◆ **19.** By comparing the results of Exercises 17 and 18, guess a solution to the **three-dimensional heat equation**

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2}.$$

Verify your guess. (If you're feeling lazy, use Maple.)

Biharmonic functions

A function $u(x, y)$ with continuous partials of fourth order is **biharmonic** if $\dfrac{\partial^2 u}{\partial x^2} + \dfrac{\partial^2 u}{\partial y^2}$ is a harmonic function.

◆ **20.** Show that $u(x, y)$ is biharmonic if and only if it satisfies the biharmonic equation

$$\frac{\partial^4 u}{\partial x^4} + 2\frac{\partial^4 u}{\partial x^2 \partial y^2} + \frac{\partial^4 u}{\partial y^4} = 0$$

21. Verify that $u(x, y) = x^4 - 3x^2y^2$ is biharmonic.

22. Show that if $u(x, y)$ is harmonic, then $v(x, y) = xu(x, y)$ and $w(x, y) = yu(x, y)$ are biharmonic.

Use the result of Exercise 22 to show that the functions in Exercises 23–25 are biharmonic.

23. $x\,e^x\sin y$ **24.** $y\ln(x^2 + y^2)$

25. $\dfrac{xy}{x^2 + y^2}$

◆ **26.** Propose a definition of a biharmonic function of three variables, and prove results analogous to those of Exercises 20 and 22 for biharmonic functions $u(x, y, z)$.

27. Use Maple to verify directly that the function of Exercise 25 is biharmonic.

12.5 The Chain Rule

The Chain Rule for functions of one variable is a formula that gives the derivative of a composition $f(g(x))$ of two functions f and g:

$$\frac{d}{dx} f(g(x)) = f'(g(x))g'(x).$$

The situation for several variables is more complicated. If f depends on more than one variable, and any of those variables can be functions of one or more other variables, we cannot expect a simple formula for partial derivatives of the composition to cover all possible cases. We must come to think of the Chain Rule as a *procedure for differentiating compositions* rather than as a formula for their derivatives. In order to motivate a formulation of the Chain Rule for functions of two variables, we begin with a concrete example.

Example 1 Suppose you are hiking in a mountainous region for which you have a map. Let (x, y) be the coordinates of your position on the map (i.e., the horizontal coordinates of your actual position in the region). Let $z = f(x, y)$ denote the height of land (above sea level, say) at position (x, y). Suppose you are walking along a trail so that your position at time t is given by $x = u(t)$ and $y = v(t)$. (These are parametric equations of the trail on the map.) At time t your altitude above sea level is given by the composite function

$$z = f(u(t), v(t)) = g(t),$$

a function of only one variable. How fast is your altitude changing with respect to time at time t?

Solution The answer is the derivative of $g(t)$:

$$\begin{aligned} g'(t) &= \lim_{h\to 0} \frac{g(t+h) - g(t)}{h} = \lim_{h\to 0} \frac{f(u(t+h), v(t+h)) - f(u(t), v(t))}{h} \\ &= \lim_{h\to 0} \frac{f(u(t+h), v(t+h)) - f(u(t), v(t+h))}{h} \\ &\qquad + \lim_{h\to 0} \frac{f(u(t), v(t+h)) - f(u(t), v(t))}{h}. \end{aligned}$$

We added 0 to the numerator of the Newton quotient in a creative way so as to separate the quotient into the sum of two quotients, in the first of which the difference of values of f involves only the first variable of f, and in the second of which the difference involves only the second variable of f. The single-variable Chain Rule suggests that the sum of the two limits above is

$$g'(t) = f_1(u(t), v(t))u'(t) + f_2(u(t), v(t))v'(t).$$

■

The above formula is the Chain Rule for $\frac{d}{dt} f(u(t), v(t))$. In terms of Leibniz notation we have

A version of the Chain Rule

If z is a function of x and y with continuous first partial derivatives, and if x and y are differentiable functions of t, then

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

Note that there are two terms in the expression for dz/dt (or $g'(t)$), one arising from each variable of f that depends on t.

Now consider a function f of two variables, x and y, each of which is in turn a function of two other variables, s and t:

$$z = f(x, y), \qquad \text{where} \qquad x = u(s, t) \quad \text{and} \quad y = v(s, t).$$

We can form the composite function

$$z = f\big(u(s, t), v(s, t)\big) = g(s, t).$$

For instance, if $f(x, y) = x^2 + 3y$, where $u(s, t) = st^2$ and $v(s, t) = s - t$, then $g(s, t) = s^2t^4 + 3(s - t)$.

Let us assume that f, u, and v have first partial derivatives with respect to their respective variables and that those of f are continuous. Then g has first partial derivatives given by

$$g_1(s, t) = f_1\big(u(s, t), v(s, t)\big)u_1(s, t) + f_2\big(u(s, t), v(s, t)\big)v_1(s, t),$$
$$g_2(s, t) = f_1\big(u(s, t), v(s, t)\big)u_2(s, t) + f_2\big(u(s, t), v(s, t)\big)v_2(s, t).$$

These formulas can be expressed more simply using Leibniz notation:

Another version of the Chain Rule

If z is a function of x and y with continuous first partial derivatives, and if x and y depend on s and t, then

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

This can be deduced from the version obtained in Example 1 by allowing u and v there to depend on two variables, but holding one of them fixed while we differentiate with respect to the other. A more formal proof of this simple but representative case of the Chain Rule will be given in the next section.

The two equations in the box above can be combined into a single matrix equation:

$$\begin{pmatrix} \dfrac{\partial z}{\partial s} & \dfrac{\partial z}{\partial t} \end{pmatrix} = \begin{pmatrix} \dfrac{\partial z}{\partial x} & \dfrac{\partial z}{\partial y} \end{pmatrix} \begin{pmatrix} \dfrac{\partial x}{\partial s} & \dfrac{\partial x}{\partial t} \\ \dfrac{\partial y}{\partial s} & \dfrac{\partial y}{\partial t} \end{pmatrix}.$$

We will comment on the significance of this matrix form at the end of the next section.

In general, if z is a function of several "primary" variables, and each of these depends on some "secondary" variables, then the partial derivative of z with respect to one of the secondary variables will have several terms, one for the contribution to the derivative arising from each of the primary variables on which z depends.

Remark Note the significance of the various subscripts denoting partial derivatives in the functional form of the Chain Rule:

$$g_1(s,t) = f_1\big(u(s,t),\, v(s,t)\big)u_1(s,t) + f_2\big(u(s,t),\, v(s,t)\big)v_1(s,t).$$

The "1" in $g_1(s,t)$ refers to differentiation with respect to s, the first variable on which g depends. By contrast, the "1" in $f_1(u(s,t), v(s,t))$ refers to differentiation with respect to x, the first variable on which f depends. (This derivative is then evaluated at $x = u(s,t)$, $y = v(s,t)$.)

Example 2 If $z = \sin(x^2 y)$, where $x = st^2$ and $y = s^2 + \dfrac{1}{t}$, find $\partial z/\partial s$ and $\partial z/\partial t$

(a) by direct substitution and the single-variable form of the Chain Rule, and

(b) by using the (two-variable) Chain Rule.

Solution

(a) By direct substitution:

$$z = \sin\left((st^2)^2\left(s^2 + \frac{1}{t}\right)\right) = \sin(s^4t^4 + s^2t^3),$$

$$\frac{\partial z}{\partial s} = (4s^3t^4 + 2st^3)\cos(s^4t^4 + s^2t^3),$$

$$\frac{\partial z}{\partial t} = (4s^4t^3 + 3s^2t^2)\cos(s^4t^4 + s^2t^3).$$

(b) Using the Chain Rule:

$$\begin{aligned}
\frac{\partial z}{\partial s} &= \frac{\partial z}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial s} \\
&= \big(2xy\cos(x^2y)\big)t^2 + \big(x^2\cos(x^2y)\big)2s \\
&= \left(2st^2\left(s^2 + \frac{1}{t}\right)t^2 + 2s^3t^4\right)\cos(s^4t^4 + s^2t^3) \\
&= (4s^3t^4 + 2st^3)\cos(s^4t^4 + s^2t^3), \\
\frac{\partial z}{\partial t} &= \frac{\partial z}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial t} \\
&= \big(2xy\cos(x^2y)\big)2st + \big(x^2\cos(x^2y)\big)\left(\frac{-1}{t^2}\right) \\
&= \left(2st^2(s^2 + \frac{1}{t})2st + s^2t^4\left(\frac{-1}{t^2}\right)\right)\cos(s^4t^4 + s^2t^3) \\
&= (4s^4t^3 + 3s^2t^2)\cos(s^4t^4 + s^2t^3).
\end{aligned}$$

Note that we still had to use direct substitution on the derivatives obtained in (b) in order to show that the values were the same as those obtained in (a).

Example 3 Find $\dfrac{\partial}{\partial x} f(x^2 y, x + 2y)$ and $\dfrac{\partial}{\partial y} f(x^2 y, x + 2y)$ in terms of the partial derivatives of f, assuming that these partial derivatives are continuous.

Solution We have

$$\begin{aligned}\frac{\partial}{\partial x} f(x^2 y, x + 2y) &= f_1(x^2 y, x + 2y)\frac{\partial}{\partial x}(x^2 y) + f_2(x^2 y, x + 2y)\frac{\partial}{\partial x}(x + 2y)\\ &= 2xy f_1(x^2 y, x + 2y) + f_2(x^2 y, x + 2y),\\ \frac{\partial}{\partial y} f(x^2 y, x + 2y) &= f_1(x^2 y, x + 2y)\frac{\partial}{\partial y}(x^2 y) + f_2(x^2 y, x + 2y)\frac{\partial}{\partial y}(x + 2y)\\ &= x^2 f_1(x^2 y, x + 2y) + 2 f_2(x^2 y, x + 2y).\end{aligned}$$

■

Example 4 Express the partial derivatives of $z = h(s, t) = f\big(g(s, t)\big)$ in terms of the derivative f' of f and the partial derivatives of g.

Solution The partial derivatives of h can be calculated using the single-variable version of the Chain Rule: if $x = g(s, t)$, then $z = f(x)$ and

$$\begin{aligned}h_1(s, t) &= \frac{\partial z}{\partial s} = \frac{dz}{dx}\frac{\partial x}{\partial s} = f'\big(g(s, t)\big) g_1(s, t),\\ h_2(s, t) &= \frac{\partial z}{\partial t} = \frac{dz}{dx}\frac{\partial x}{\partial t} = f'\big(g(s, t)\big) g_2(s, t).\end{aligned}$$

■

The following example involves a hybrid application of the Chain Rule to a function that depends both directly and indirectly on the variable of differentiation.

Example 5 Find dz/dt, where $z = f(x, y, t)$, $x = g(t)$, and $y = h(t)$. (Assume that f, g, and h all have continuous derivatives.)

Solution Since z depends on t through each of the three variables of f, there will be three terms in the appropriate Chain Rule:

$$\begin{aligned}\frac{dz}{dt} &= \frac{\partial z}{\partial x}\frac{dx}{dt} + \frac{\partial z}{\partial y}\frac{dy}{dt} + \frac{\partial z}{\partial t}\\ &= f_1(x, y, t) g'(t) + f_2(x, y, t) h'(t) + f_3(x, y, t).\end{aligned}$$

■

Remark In the above example we can easily distinguish between the meanings of the symbols dz/dt and $\partial z/\partial t$. If, however, we had been dealing with the situation

$$z = f(x, y, s, t), \qquad \text{where} \quad x = g(s, t) \quad \text{and} \quad y = h(s, t),$$

then the meaning of the symbol $\partial z/\partial t$ would be unclear; it could refer to the simple partial derivative of f with respect to its fourth primary variable (i.e., $f_4(x, y, s, t)$), or it could refer to the derivative of the composite function $f(g(s, t), h(s, t), s, t)$. Three of the four primary variables of f depend on t and, therefore, contribute to the rate of change of z with respect to t. The partial derivative $f_4(x, y, s, t)$ denotes the contribution of only one of these three variables. It is conventional to use $\partial z/\partial t$

to denote the derivative of the composite function with respect to the secondary variable t:

$$\begin{aligned}\frac{\partial z}{\partial t} &= \frac{\partial}{\partial t} f(g(s,t), h(s,t), s, t) \\ &= f_1(x, y, s, t) g_2(s,t) + f_2(x, y, s, t) h_2(s,t) + f_4(x, y, s, t).\end{aligned}$$

When it is necessary, we can denote the contribution coming from the primary variable t by

$$\left(\frac{\partial z}{\partial t}\right)_{x,y,s} = \frac{\partial}{\partial t} f(x, y, s, t) = f_4(x, y, s, t).$$

Here, the subscripts denote those primary variables of f whose contributions to the rate of change of z with respect to t are being *ignored.* Of course, in the situation described above, $(\partial z/\partial t)_s$ means the same as $\partial z/\partial t$.

In applications, the variables that contribute to a particular partial derivative will usually be clear from the context. The following example contains such an application. This is an example of a procedure called *differentiation following the motion.*

Example 6 Atmospheric temperature depends on position and time. If we denote position by three spatial coordinates x, y, and z (measured in kilometres), and time by t (measured in hours), then the temperature T is a function of four variables, $T(x, y, z, t)$.

(a) If a thermometer is attached to a weather balloon that moves through the atmosphere on a path with parametric equations $x = f(t)$, $y = g(t)$, and $z = h(t)$, what is the rate of change at time T of the temperature recorded by the thermometer?

(b) Find the rate of change of the recorded temperature at time $t = 1$ if

$$T(x, y, z, t) = \frac{xy}{1+z}(1+t),$$

and if the balloon moves along the curve

$$x = t, \qquad y = 2t, \qquad z = t - t^2.$$

Solution

(a) Here, the rate of change of the thermometer reading depends on the change in position of the thermometer as well as increasing time. Thus, none of the four variables of T can be ignored in the differentiation. The rate is given by

$$\frac{dT}{dt} = \frac{\partial T}{\partial x}\frac{dx}{dt} + \frac{\partial T}{\partial y}\frac{dy}{dt} + \frac{\partial T}{\partial z}\frac{dz}{dt} + \frac{\partial T}{\partial t}.$$

The term $\partial T/\partial t$ refers only to the rate of change of the temperature with respect to time at a fixed position in the atmosphere. The other three terms arise from the motion of the balloon.

(b) The values of the three coordinates and their derivatives at $t = 1$ are $x = 1$, $y = 2$, $z = 0$, $dx/dt = 1$, $dy/dt = 2$, and $dz/dt = -1$. Also, at $t = 1$,

$$\frac{\partial T}{\partial x} = \frac{y}{1+z}(1+t) = 4, \qquad \frac{\partial T}{\partial z} = \frac{-xy}{(1+z)^2}(1+t) = -4,$$
$$\frac{\partial T}{\partial y} = \frac{x}{1+z}(1+t) = 2, \qquad \frac{\partial T}{\partial t} = \frac{xy}{1+z} = 2.$$

Thus,

$$\left.\frac{dT}{dt}\right|_{t=1} = (4)(1) + (2)(2) + (-4)(-1) + 2 = 14.$$

The recorded temperature is increasing at a rate of 14 °/h at time $t = 1$. ■

The discussion and examples above show that the Chain Rule for functions of several variables can take different forms depending on the numbers of variables of the various functions being composed. As an aid in determining the correct form of the Chain Rule in a given situation you can construct a chart showing which variables depend on which. Figure 12.21(a) shows such a chart for the temperature function of Example 6. The Chain Rule for dT/dt involves a term for every route from T to t in the chart. The route from T through x to t produces the term $\frac{\partial T}{\partial x}\frac{dx}{dt}$ and so on.

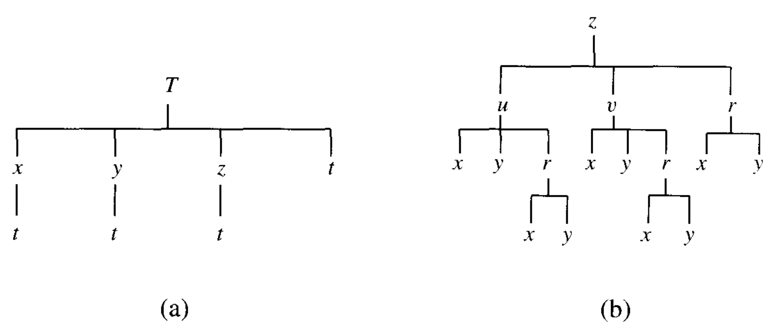

Figure 12.21

(a) Chart showing the dependence of T on t in Example 6

(b) Dependence chart for Example 7

Example 7 Write the appropriate Chain Rule for $\partial z/\partial x$, where z depends on u, v, and r; u and v depend on x, y, and r; and r depends on x and y.

Solution The appropriate chart is shown in Figure 12.21(b). There are five routes from z to x:

$$\frac{\partial z}{\partial x} = \frac{\partial z}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial z}{\partial u}\frac{\partial u}{\partial r}\frac{\partial r}{\partial x} + \frac{\partial z}{\partial v}\frac{\partial v}{\partial x} + \frac{\partial z}{\partial v}\frac{\partial v}{\partial r}\frac{\partial r}{\partial x} + \frac{\partial z}{\partial r}\frac{\partial r}{\partial x}.$$

■

Homogeneous Functions

A function $f(x_1, \ldots, x_n)$ is said to be **positively homogeneous of degree** k if, for every point $(x_1, x_2, \ldots, x_n)$ in its domain and every real number $t > 0$, we have

$$f(tx_1, tx_2, \ldots, tx_n) = t^k f(x_1, \ldots, x_n).$$

For example,

$$f(x, y) = x^2 + xy - y^2 \quad \text{is positively homogeneous of degree 2,}$$
$$f(x, y) = \sqrt{x^2 + y^2} \quad \text{is positively homogeneous of degree 1,}$$
$$f(x, y) = \frac{2xy}{x^2 + y^2} \quad \text{is positively homogeneous of degree 0,}$$
$$f(x, y, z) = \frac{x - y + 5z}{yz - z^2} \quad \text{is positively homogeneous of degree } -1,$$
$$f(x, y) = x^2 + y \quad \text{is not positively homogeneous.}$$

Observe that a positively homogeneous function of degree 0 remains constant along rays from the origin. More generally, along such rays a positively homogeneous function of degree k grows or decays proportionally to the kth power of distance from the origin.

THEOREM 2

Euler's Theorem

If $f(x_1, \ldots, x_n)$ has continuous first partial derivatives and is positively homogeneous of degree k, then

$$\sum_{i=1}^{n} x_i f_i(x_1, \ldots, x_n) = kf(x_1, \ldots, x_n).$$

PROOF Differentiate the equation $f(tx_1, tx_2, \ldots, tx_n) = t^k f(x_1, \ldots, x_n)$ with respect to t to get

$$x_1 f_1(tx_1, \ldots, tx_n) + x_2 f_2(tx_1, \ldots, tx_n) + \ldots + x_n f_n(tx_1, \ldots, tx_n) = k\,t^{k-1} f(x_1, \ldots, x_n).$$

Now substitute $t = 1$ to get the desired result.

Note that Exercises 26–29 in Section 12.3 illustrate this theorem.

Higher-Order Derivatives

Applications of the Chain Rule to higher-order derivatives can become quite complicated. It is important to keep in mind at each stage which variables are independent of one another.

Example 8 Calculate $\dfrac{\partial^2}{\partial x \partial y} f(x^2 - y^2, xy)$ in terms of partial derivatives of the function f. Assume that the second-order partials of f are continuous.

Solution In this problem symbols for the primary variables on which f depends are not stated explicitly. Let them be u and v. The problem therefore asks us to find

$$\frac{\partial^2}{\partial x \partial y} f(u,v), \qquad \text{where} \quad u = x^2 - y^2 \quad \text{and} \quad v = xy.$$

First differentiate with respect to y:

$$\frac{\partial}{\partial y} f(u,v) = -2y f_1(u,v) + x f_2(u,v).$$

Now differentiate this result with respect to x. Note that the second term on the right is a product of two functions of x so we need to use the Product Rule:

$$\begin{aligned}\frac{\partial^2}{\partial x \partial y} f(u,v) = & -2y\Big(2x f_{11}(u,v) + y f_{12}(u,v)\Big)\\ & + f_2(u,v) + x\Big(2x f_{21}(u,v) + y f_{22}(u,v)\Big)\\ = & f_2(u,v) - 4xy f_{11}(u,v) + 2(x^2 - y^2) f_{12}(u,v) + xy f_{22}(u,v).\end{aligned}$$

In the last step we have used the fact that the mixed partials of f are continuous so we could equate f_{12} and f_{21}.

Review the above calculation very carefully and make sure you understand what is being done at each step. Note that all the derivatives of f that appear are evaluated at $(u,v) = (x^2 - y^2, xy)$, not at (x,y), because x and y are not themselves the primary variables on which f depends.

Remark The kind of calculation done in the above example (and the following ones) is easily carried out by a computer algebra system. In Maple:

```
> g := (x,y) -> f(x^2 - y^2, x*y);
   simplify(D[1,2](g)(x,y));
```

$$\begin{aligned}&-4y D_{1,1}(f)(x^2 - y^2, xy)x - 2 * D_{1,2}(f)(x^2 - y^2, xy)y^2\\ &\quad + 2D_{1,2}(f)(x^2 - y^2, xy)x^2\\ &\quad + x D_{2,2}(f)(x^2 - y^2, xy)y + D_2(f)(x^2 - y^2, xy)\end{aligned}$$

which, on close inspection, is the same answer we calculated in the example.

Example 9 Show that $f(x^2 - y^2, 2xy)$ is a harmonic function if $f(x,y)$ is harmonic.

Solution Let $u = x^2 - y^2$ and $v = 2xy$. If $z = f(u,v)$, then

$$\begin{aligned}\frac{\partial z}{\partial x} &= 2x f_1(u,v) + 2y f_2(u,v),\\ \frac{\partial z}{\partial y} &= -2y f_1(u,v) + 2x f_2(u,v),\\ \frac{\partial^2 z}{\partial x^2} &= 2f_1(u,v) + 2x\big(2x f_{11}(u,v) + 2y f_{12}(u,v)\big)\\ &\quad + 2y\big(2x f_{21}(u,v) + 2y f_{22}(u,v)\big)\\ &= 2f_1(u,v) + 4x^2 f_{11}(u,v) + 8xy f_{12}(u,v) + 4y^2 f_{22}(u,v),\end{aligned}$$

$$\begin{aligned}\frac{\partial^2 z}{\partial y^2} &= -2f_1(u,v) - 2y\big(-2yf_{11}(u,v) + 2xf_{12}(u,v)\big)\\ &\quad + 2x\big(-2yf_{21}(u,v) + 2xf_{22}(u,v)\big)\\ &= -2f_1(u,v) + 4y^2 f_{11}(u,v) - 8xyf_{12}(u,v) + 4x^2 f_{22}(u,v).\end{aligned}$$

Therefore,

$$\frac{\partial^2 z}{\partial x^2} + \frac{\partial^2 z}{\partial y^2} = 4(x^2+y^2)\big(f_{11}(u,v) + f_{22}(u,v)\big) = 0$$

because f is given to be harmonic. Thus, $z = f(x^2 - y^2, 2xy)$ is a harmonic function of x and y.

In the following example we show that the two-dimensional Laplace differential equation (see Example 3 in Section 12.4) takes the form

$$\frac{\partial^2 z}{\partial r^2} + \frac{1}{r}\frac{\partial z}{\partial r} + \frac{1}{r^2}\frac{\partial^2 z}{\partial \theta^2} = 0$$

when stated for a function z expressed in terms of polar coordinates r and θ.

Example 10 **(Laplace's equation in polar coordinates)** If $z = f(x, y)$ has continuous partial derivatives of second order, and if $x = r\cos\theta$ and $y = r\sin\theta$, show that

$$\frac{\partial^2 z}{\partial r^2} + \frac{1}{r}\frac{\partial z}{\partial r} + \frac{1}{r^2}\frac{\partial^2 z}{\partial \theta^2} = \frac{\partial^2 z}{\partial x^2} + \frac{\partial^2 z}{\partial y^2}.$$

Solution It is possible to do this in two different ways; we can start with either side and use the Chain Rule to show that it is equal to the other side. Here, we will calculate the partial derivatives with respect to r and θ that appear on the left side and express them in terms of partial derivatives with respect to x and y. The other approach, involving expressing partial derivatives with respect to x and y in terms of partial derivatives with respect to r and θ, is a little harder. (See Exercise 24 below.) However, we would have to do it that way if we were not given the form of the differential equation in polar coordinates and had to find it.

First note that

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

Thus,

$$\frac{\partial z}{\partial r} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial r} = \cos\theta\,\frac{\partial z}{\partial x} + \sin\theta\,\frac{\partial z}{\partial y}.$$

BEWARE
This is a difficult but important example. Examine each step carefully to make sure you understand what is being done.

Now differentiate with respect to r again. Remember that r and θ are independent variables, so the factors $\cos\theta$ and $\sin\theta$ can be regarded as constants. However, $\partial z/\partial x$ and $\partial z/\partial y$ depend on x and y and, therefore, on r and θ.

$$\begin{aligned}\frac{\partial^2 z}{\partial r^2} &= \cos\theta \frac{\partial}{\partial r}\frac{\partial z}{\partial x} + \sin\theta \frac{\partial}{\partial r}\frac{\partial z}{\partial y}\\ &= \cos\theta\left(\cos\theta \frac{\partial^2 z}{\partial x^2} + \sin\theta \frac{\partial^2 z}{\partial y \partial x}\right) + \sin\theta\left(\cos\theta \frac{\partial^2 z}{\partial x \partial y} + \sin\theta \frac{\partial^2 z}{\partial y^2}\right)\\ &= \cos^2\theta \frac{\partial^2 z}{\partial x^2} + 2\cos\theta\sin\theta \frac{\partial^2 z}{\partial x \partial y} + \sin^2\theta \frac{\partial^2 z}{\partial y^2}.\end{aligned}$$

We have used the equality of mixed partials in the last line. Similarly,

$$\frac{\partial z}{\partial \theta} = -r\sin\theta \frac{\partial z}{\partial x} + r\cos\theta \frac{\partial z}{\partial y}.$$

When we differentiate a second time with respect to θ we can regard r as constant, but each term above is still a product of two functions that depend on θ. Thus,

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

Combining these results, we obtain the desired formula:

$$\frac{\partial^2 z}{\partial r^2} + \frac{1}{r}\frac{\partial z}{\partial r} + \frac{1}{r^2}\frac{\partial^2 z}{\partial \theta^2} = \frac{\partial^2 z}{\partial x^2} + \frac{\partial^2 z}{\partial y^2}.$$

Exercises 12.5

In Exercises 1–4, write appropriate versions of the Chain Rule for the indicated derivatives.

1. $\partial w/\partial t$ if $w = f(x, y, z)$, where $x = g(s, t)$, $y = h(s, t)$, and $z = k(s, t)$
2. $\partial w/\partial t$ if $w = f(x, y, z)$, where $x = g(s)$, $y = h(s, t)$, and $z = k(t)$
3. $\partial z/\partial u$ if $z = g(x, y)$, where $y = f(x)$ and $x = h(u, v)$
4. dw/dt if $w = f(x, y)$, $x = g(r, s)$, $y = h(r, t)$, $r = k(s, t)$, and $s = m(t)$
5. If $w = f(x, y, z)$, where $x = g(y, z)$ and $y = h(z)$, state appropriate versions of the Chain Rule for $\dfrac{dw}{dz}$, $\left(\dfrac{\partial w}{\partial z}\right)_x$, and $\left(\dfrac{\partial w}{\partial z}\right)_{x,y}$.
6. Use two different methods to calculate $\partial u/\partial t$ if $u = \sqrt{x^2 + y^2}$, $x = e^{st}$, and $y = 1 + s^2\cos t$.
7. Use two different methods to calculate $\partial z/\partial x$ if $z = \tan^{-1}(u/v)$, $u = 2x + y$, and $v = 3x - y$.
8. Use two methods to calculate dz/dt given that $z = txy^2$, $x = t + \ln(y + t^2)$, and $y = e^t$.

In Exercises 9–12, find the indicated derivatives, assuming that the function $f(x, y)$ has continuous first partial derivatives.

9. $\dfrac{\partial}{\partial x} f(2x, 3y)$
10. $\dfrac{\partial}{\partial x} f(2y, 3x)$
11. $\dfrac{\partial}{\partial x} f(y^2, x^2)$
12. $\dfrac{\partial}{\partial y} f\big(yf(x, t), f(y, t)\big)$
13. Suppose that the temperature T in a certain liquid varies with depth z and time t according to the formula $T = e^{-t}z$. Find the rate of change of temperature with respect to time

at a point that is moving through the liquid so that at time t its depth is $f(t)$. What is this rate if $f(t) = e^t$? What is happening in this case?

14. Suppose the strength E of an electric field in space varies with position (x, y, z) and time t according to the formula $E = f(x, y, z, t)$. Find the rate of change with respect to time of the electric field strength measured by an instrument moving through space along the helix $x = \sin t$, $y = \cos t$, $z = t$.

In Exercises 15–20, assume that f has continuous partial derivatives of all orders.

15. If $z = f(x, y)$, where $x = 2s + 3t$ and $y = 3s - 2t$, find

(a) $\dfrac{\partial^2 z}{\partial s^2}$, (b) $\dfrac{\partial^2 z}{\partial s \partial t}$, and (c) $\dfrac{\partial^2 z}{\partial t^2}$.

16. If $f(x, y)$ is harmonic, show that $f\left(\dfrac{x}{x^2+y^2}, \dfrac{-y}{x^2+y^2}\right)$ is also harmonic.

17. If $x = t \sin s$ and $y = t \cos s$, find $\dfrac{\partial^2}{\partial s \partial t} f(x, y)$.

18. Find $\dfrac{\partial^3}{\partial x \partial y^2} f(2x + 3y, xy)$ in terms of partial derivatives of f.

19. Find $\dfrac{\partial^2}{\partial y \partial x} f(y^2, xy, -x^2)$ in terms of partial derivatives of the function f.

20. Find $\dfrac{\partial^3}{\partial t^2 \partial s} f(s^2 - t, s + t^2)$ in terms of partial derivatives of f.

21. Suppose that $u(x, y)$ and $v(x, y)$ have continuous second partial derivatives and satisfy the Cauchy–Riemann equations

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \quad \text{and} \quad \frac{\partial v}{\partial x} = -\frac{\partial u}{\partial y}.$$

Suppose also that $f(u, v)$ is a harmonic function of u and v. Show that $f\big(u(x, y), v(x, y)\big)$ is a harmonic function of x and y. *Hint:* u and v are harmonic functions by Exercise 15 in Section 12.4.

22. If $r^2 = x^2 + y^2 + z^2$, verify that $u(x, y, z) = 1/r$ is harmonic throughout $\mathbb{R}^3$ except at the origin.

* **23.** If $x = e^s \cos t$, $y = e^s \sin t$, and $z = u(x, y) = v(s, t)$, show that

$$\frac{\partial^2 z}{\partial s^2} + \frac{\partial^2 z}{\partial t^2} = (x^2 + y^2)\left(\frac{\partial^2 z}{\partial x^2} + \frac{\partial^2 z}{\partial y^2}\right).$$

* **24. (Converting Laplace's equation to polar coordinates)** The transformation to polar coordinates, $x = r \cos\theta$, $y = r \sin\theta$, implies that $r^2 = x^2 + y^2$ and $\tan\theta = y/x$. Use these equations to show that

$$\frac{\partial r}{\partial x} = \cos\theta \qquad \frac{\partial r}{\partial y} = \sin\theta$$
$$\frac{\partial \theta}{\partial x} = -\frac{\sin\theta}{r} \qquad \frac{\partial \theta}{\partial y} = \frac{\cos\theta}{r}.$$

Use these formulas to help you express $\dfrac{\partial^2 u}{\partial x^2} + \dfrac{\partial^2 u}{\partial y^2}$ in terms of partials of u with respect to r and θ, and hence reprove the formula for the Laplace differential equation in polar coordinates given in Example 10.

25. If $u(x, y) = r^2 \ln r$, where $r^2 = x^2 + y^2$, verify that u is a biharmonic function by showing that

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right)\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}\right) = 0.$$

26. If $f(x, y)$ is positively homogeneous of degree k and has continuous partial derivatives of second order, show that

$$x^2 f_{11}(x, y) + 2xy f_{12}(x, y) + y^2 f_{22}(x, y) \\ = k(k-1) f(x, y).$$

* **27.** Generalize the result of Exercise 26 to functions of n variables.

* **28.** Generalize the results of Exercises 26 and 27 to expressions involving mth-order partial derivatives of the function f.

Exercises 29–30 revisit Exercise 16 of Section 12.4. Let

$$F(x, y) = \begin{cases} \dfrac{2xy(x^2 - y^2)}{x^2 + y^2}, & \text{if } (x, y) \neq (0, 0) \\ 0, & \text{if } (x, y) = (0, 0). \end{cases}$$

29. (a) Show that $F(x, y) = -F(y, x)$ for all (x, y).

(b) Show that $F_1(x, y) = -F_2(y, x)$ and $F_{12}(x, y) = -F_{21}(y, x)$ for $(x, y) \neq (0, 0)$.

(c) Show that $F_1(0, y) = -2y$ for all y and, hence, that $F_{12}(0, 0) = 2$.

(d) Deduce that $F_2(x, 0) = 2x$ and $F_{21}(0, 0) = 2$.

30. (a) Use Exercise 29(b) to find $F_{12}(x, x)$ for $x \neq 0$.

(b) Is $F_{12}(x, y)$ continuous at $(0, 0)$? Why?

◆ **31.** Use the change of variables $\xi = x + ct$, $\eta = x$ to transform the partial differential equation

$$\frac{\partial u}{\partial t} = c\,\frac{\partial u}{\partial x}, \qquad (c = \text{constant}),$$

into the simpler equation $\partial v/\partial \eta = 0$, where $v(\xi, \eta) = v(x + ct, x) = u(x, t)$. This equation says that $v(\xi, \eta)$ does not depend on η, so $v = f(\xi)$ for some arbitrary differentiable function f. What is the corresponding "general solution" $u(x, t)$ of the original partial differential equation?

◈ **32.** Having considered Exercise 31, guess a "general solution" $w(r, s)$ of the second-order partial differential equation

$$\frac{\partial^2}{\partial r\,\partial s}\,w(r, s) = 0.$$

Your answer should involve two arbitrary functions.

◈ **33.** Use the change of variables $r = x + ct$, $s = x - ct$, $w(r, s) = u(x, t)$ to transform the one-dimensional wave equation

$$\frac{\partial^2 u}{\partial t^2} = c^2\,\frac{\partial^2 u}{\partial x^2}$$

to a simpler form. Now use the result of Exercise 32 to find the *general solution* of this wave equation in the form given in Example 4 in Section 12.4.

◈ **34.** Show that the initial-value problem for the one-dimensional wave equation

$$\begin{cases} u_{tt}(x, t) = c^2 u_{xx}(x, t) \\ u(x, 0) = p(x) \\ u_t(x, 0) = q(x) \end{cases}$$

has the solution

$$u(x, t) = \frac{1}{2}\Big[p(x-ct)+p(x+ct)\Big]+\frac{1}{2c}\int_{x-ct}^{x+ct} q(s)\,ds.$$

(Note that we have used subscripts "x" and "t" instead of "1" and "2" to denote the partial derivatives here. This is common usage in dealing with partial differential equations.)

Remark The initial-value problem in Exercise 33 gives the small lateral displacement $u(x, t)$ at position x at time t of a vibrating string held under tension along the x-axis. The function $p(x)$ gives the *initial* displacement at position x, that is, the displacement at time $t = 0$. Similarly, $q(x)$ gives the initial velocity at position x. Observe that the position at time t depends only on values of these initial data at points no further than ct units away. This is consistent with the previous observation that the solutions of the wave equation represent waves travelling with speed c.

Redo the Examples and Exercises listed in Exercises 35–40 using Maple to do the calculations.

35. Example 10

36. Exercise 16

37. Exercise 19

38. Exercise 20

39. Exercise 23

40. Exercise 34

12.6 Linear Approximations, Differentiability, and Differentials

The tangent line to the graph $y = f(x)$ at $x = a$ provides a convenient approximation for values of $f(x)$ for x near a (see Figure 12.22:

$$f(x) \approx L(x) = f(a) + f'(a)(x - a).$$

Here, $L(x)$ is the **linearization** of f at a; its graph is the tangent line to $y = f(x)$ there. The mere existence of $f'(a)$ is sufficient to guarantee that the error in the approximation (the vertical distance between the curve and tangent at x) is small compared with the distance $h = x - a$ between a and x, that is,

$$\begin{aligned}\lim_{h\to 0}\frac{f(a+h) - L(a+h)}{h} &= \lim_{h\to 0}\frac{f(a+h) - f(a) - f'(a)h}{h} \\ &= \lim_{h\to 0}\frac{f(a+h) - f(a)}{h} - f'(a) \\ &= f'(a) - f'(a) = 0.\end{aligned}$$

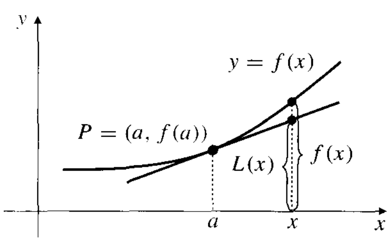

Figure 12.22 The linearization of f at $x = a$

Similarly, the tangent plane to the graph of $z = f(x, y)$ at (a, b) is $z = L(x, y)$, where

$$L(x, y) = f(a, b) + f_1(a, b)(x - a) + f_2(a, b)(y - b)$$

is the **linearization** of f at (a, b). We can use $L(x, y)$ to approximate values of $f(x, y)$ near (a, b):

$$f(x, y) \approx L(x, y) = f(a, b) + f_1(a, b)(x - a) + f_2(a, b)(y - b).$$

Example 1 Find an approximate value for $f(x, y) = \sqrt{2x^2 + e^{2y}}$ at $(2.2, -0.2)$.

Solution It is convenient to use the linearization at $(2, 0)$, where the values of f and its partials are easily evaluated:

$$\begin{aligned} & & f(2,0) &= 3, \\ f_1(x,y) &= \frac{2x}{\sqrt{2x^2+e^{2y}}}, & f_1(2,0) &= \frac{4}{3}, \\ f_2(x,y) &= \frac{e^{2y}}{\sqrt{2x^2+e^{2y}}}, & f_2(2,0) &= \frac{1}{3}. \end{aligned}$$

Thus, $L(x, y) = 3 + \frac{4}{3}(x - 2) + \frac{1}{3}(y - 0)$, and

$$f(2.2, -0.2) \approx L(2.2, -0.2) = 3 + \frac{4}{3}(2.2 - 2) + \frac{1}{3}(-0.2 - 0) = 3.2\,.$$

(For the sake of comparison, $f(2.2, -0.2) \approx 3.2172$ to 4 decimal places.)

Unlike the single-variable case, the mere existence of the partial derivatives $f_1(a, b)$ and $f_2(a, b)$ does not even imply that f is continuous at (a, b), let alone that the error in the linearization is small compared with the distance $\sqrt{(x-a)^2 + (y-b)^2}$ between (a, b) and (x, y). We adopt this latter condition as our definition of what it means for a function of two variables to be *differentiable* at a point.

DEFINITION 5

We say that the function $f(x, y)$ is **differentiable** at the point (a, b) if

$$\lim_{(h,k)\to(0,0)} \frac{f(a+h, b+k) - f(a, b) - h\,f_1(a, b) - kf_2(a, b)}{\sqrt{h^2 + k^2}} = 0.$$

This definition and the following theorems can be generalized to functions of any number of variables in the obvious way. For the sake of simplicity, we state them for the two-variable case only.

The function $f(x, y)$ is differentiable at the point (a, b) if and only if the surface $z = f(x, y)$ has a *nonvertical tangent plane* at (a, b). This implies that $f_1(a, b)$ and $f_2(a, b)$ must exist and that f must be continuous at (a, b). (Recall, however, that the existence of the partial derivatives does *not* even imply that f is continuous, let alone differentiable.) In particular, the function is *continuous* wherever it is differentiable. We will prove a two-variable version of the Mean-Value Theorem and use it to show that functions are differentiable wherever they have *continuous* first partial derivatives.

THEOREM 3

A Mean-Value Theorem

If $f_1(x, y)$ and $f_2(x, y)$ are continuous in a neighbourhood of the point (a, b), and if the absolute values of h and k are sufficiently small, then there exist numbers θ_1 and θ_2, each between 0 and 1, such that

$$f(a + h, b + k) - f(a, b) = hf_1(a + \theta_1 h, b + k) + kf_2(a, b + \theta_2 k).$$

PROOF The proof of this theorem is very similar to that of Theorem 1 in Section 12.4, so we give only a sketch here. The reader can fill in the details. Write

$$f(a+h, b+k) - f(a, b) = \big(f(a+h, b+k) - f(a, b+k)\big) + \big(f(a, b+k) - f(a, b)\big),$$

and then apply the single-variable Mean-Value Theorem separately to $f(x, b+k)$ on the interval between a and $a+h$, and to $f(a, y)$ on the interval between b and $b+k$ to get the desired result.

THEOREM 4

If f_1 and f_2 are continuous in a neighbourhood of the point (a, b), then f is differentiable at (a, b).

PROOF Using Theorem 3 and the facts that

$$\left|\frac{h}{\sqrt{h^2+k^2}}\right| \le 1 \qquad \text{and} \qquad \left|\frac{k}{\sqrt{h^2+k^2}}\right| \le 1,$$

we estimate

$$\begin{aligned}
&\left|\frac{f(a+h, b+k) - f(a, b) - hf_1(a, b) - kf_2(a, b)}{\sqrt{h^2+k^2}}\right| \\
&\quad = \left|\frac{h}{\sqrt{h^2+k^2}}\big(f_1(a+\theta_1 h, b+k) - f_1(a, b)\big)\right. \\
&\qquad \left. + \frac{k}{\sqrt{h^2+k^2}}\big(f_2(a, b+\theta_2 k) - f_2(a, b)\big)\right| \\
&\quad \le \big|f_1(a+\theta_1 h, b+k) - f_1(a, b)\big| + \big|f_2(a, b+\theta_2 k) - f_2(a, b)\big|.
\end{aligned}$$

Since f_1 and f_2 are continuous at (a, b), each of these latter terms approaches 0 as h and k approach 0. This is what we needed to prove.

We illustrate differentiability with an example where we can calculate directly the error in the tangent plane approximation.

Example 2 Calculate $f(x+h, y+k) - f(x, y) - f_1(x, y)h - f_2(x, y)k$ if $f(x, y) = x^3 + xy^2$.

Solution Since $f_1(x, y) = 3x^2 + y^2$ and $f_2(x, y) = 2xy$, we have

$$\begin{aligned}
&f(x+h, y+k) - f(x, y) - f_1(x, y)h - f_2(x, y)k \\
&\qquad = (x+h)^3 + (x+h)(y+k)^2 - x^3 - xy^2 - (3x^2+y^2)h - 2xyk \\
&\qquad = 3xh^2 + h^3 + 2yhk + hk^2 + xk^2.
\end{aligned}$$

Observe that the result above is a polynomial in h and k with no term of degree less than 2 in these variables. Therefore, this difference approaches zero like the *square* of the distance $\sqrt{h^2+k^2}$ from (x, y) to $(x+h, y+k)$ as $(h, k) \to (0, 0)$, so the condition for differentiability is certainly satisfied:

$$\lim_{(h,k)\to(0,0)} \frac{3xh^2 + h^3 + 2yhk + hk^2 + xk^2}{\sqrt{h^2+k^2}} = 0.$$

This quadratic behaviour is the case for any function f with continuous *second* partial derivatives. (See Exercise 19 below.)

Proof of the Chain Rule

We are now able to give a formal statement and proof of a simple but representative case of the Chain Rule for multivariate functions.

THEOREM 5

A Chain Rule

Let $z = f(x, y)$, where $x = u(s, t)$ and $y = v(s, t)$. Suppose that

(i) $u(a, b) = p$ and $v(a, b) = q$,

$w_1(a, b) = f_1(p, q)u_1(a, b) + f_2(p, q)v_1(a, b)$,

$w_2(a, b) = f_1(p, q)u_2(a, b) + f_2(p, q)v_2(a, b)$,

that is,

$$\frac{\partial z}{\partial s} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial s} \quad \text{and} \quad \frac{\partial z}{\partial t} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial t}.$$

PROOF Define a function E of two variables as follows: $E(0, 0) = 0$, and if $(h, k) \neq (0, 0)$, then

$$E(h, k) = \frac{f(p + h, q + k) - f(p, q) - hf_1(p, q) - kf_2(p, q)}{\sqrt{h^2 + k^2}}.$$

Observe that $E(h, k)$ is continuous at $(0, 0)$ because f is differentiable at (p, q). Now,

$$f(p + h, q + k) - f(p, q) = hf_1(p, q) + kf_2(p, q) + \sqrt{h^2 + k^2}\, E(h, k).$$

In this formula put $h = u(a + \sigma, b) - u(a, b)$ and $k = v(a + \sigma, b) - v(a, b)$ and divide by σ to obtain

$$\begin{aligned}\frac{w(a + \sigma, b) - w(a, b)}{\sigma} &= \frac{f(u(a + \sigma, b), v(a + \sigma, b)) - f(u(a, b), v(a, b))}{\sigma} \\ &= \frac{f(p + h, q + k) - f(p, q)}{\sigma} \\ &= f_1(p, q)\frac{h}{\sigma} + f_2(p, q)\frac{k}{\sigma} + \sqrt{\left(\frac{h}{\sigma}\right)^2 + \left(\frac{k}{\sigma}\right)^2}\, E(h, k).\end{aligned}$$

We want to let σ approach 0 in this formula. Note that

$$\lim_{\sigma \to 0} \frac{h}{\sigma} = \lim_{\sigma \to 0} \frac{u(a + \sigma, b) - u(a, b)}{\sigma} = u_1(a, b),$$

and, similarly, $\lim_{\sigma \to 0}(k/\sigma) = v_1(a, b)$. Since $(h, k) \to (0, 0)$ if $\sigma \to 0$, we have

$$w_1(a, b) = f_1(p, q)u_1(a, b) + f_2(p, q)v_1(a, b).$$

The proof for w_2 is similar.

Differentials

If the first partial derivatives of a function $z = f(x_1, \ldots, x_n)$ exist at a point, we may construct a **differential** dz or df of the function at that point in a manner similar to that used for functions of one variable:

$$\begin{aligned} dz = df &= \frac{\partial z}{\partial x_1}\,dx_1 + \frac{\partial z}{\partial x_2}\,dx_2 + \cdots + \frac{\partial z}{\partial x_n}\,dx_n \\ &= f_1(x_1, \ldots, x_n)\,dx_1 + \cdots + f_n(x_1, \ldots, x_n)\,dx_n. \end{aligned}$$

Here, the differential dz is considered to be a function of the $2n$ independent variables $x_1, x_2, \ldots, x_n, dx_1, dx_2, \ldots, dx_n$.

For a differentiable function f, the differential df is an approximation to the change Δf in value of the function given by

$$\Delta f = f(x_1 + dx_1, \ldots, x_n + dx_n) - f(x_1, \ldots, x_n).$$

The error in this approximation is small compared with the distance between the two points in the domain of f, that is,

$$\frac{\Delta f - df}{\sqrt{(dx_1)^2 + \cdots + (dx_n)^2}} \to 0 \quad \text{if all } dx_i \to 0, \qquad (1 \le i \le n).$$

In this sense, differentials are just another way of looking at linearization.

Example 3 Estimate the percentage change in the period

$$T = 2\pi\sqrt{\frac{L}{g}}$$

of a simple pendulum if the length L of the pendulum increases by 2% and the acceleration of gravity g decreases by 0.6%.

Solution We calculate the differential of T:

$$\begin{aligned} dT &= \frac{\partial T}{\partial L}\,dL + \frac{\partial T}{\partial g}\,dg \\ &= \frac{2\pi}{2\sqrt{Lg}}\,dL - \frac{2\pi\sqrt{L}}{2g^{3/2}}\,dg. \end{aligned}$$

We are given that $dL = \dfrac{2}{100}L$ and $dg = -\dfrac{6}{1{,}000}g$. Thus

$$dT = \frac{1}{100}2\pi\sqrt{\frac{L}{g}} - \left(-\frac{6}{1{,}000}\right)\frac{2\pi}{2}\sqrt{\frac{L}{g}} = \frac{13}{1{,}000}T.$$

Therefore, the period T of the pendulum increases by 1.3%. ■

Functions from n-space to m-space

(This is an optional topic.) A vector $\mathbf{f} = (f_1, f_2, \ldots, f_m)$ of m functions, each depending on n variables $(x_1, x_2, \ldots, x_n)$, defines a *transformation* (i.e., a function) from $\mathbb{R}^n$ to $\mathbb{R}^m$; specifically, if $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ is a point in $\mathbb{R}^n$, and

$$\begin{aligned} y_1 &= f_1(x_1, x_2, \ldots, x_n) \\ y_2 &= f_2(x_1, x_2, \ldots, x_n) \\ &\vdots \\ y_m &= f_m(x_1, x_2, \ldots, x_n), \end{aligned}$$

then $\mathbf{y} = (y_1, y_2, \ldots, y_m)$ is the point in $\mathbb{R}^m$ that corresponds to $\mathbf{x}$ under the transformation $\mathbf{f}$. We can write these equations more compactly as

$$\mathbf{y} = \mathbf{f}(\mathbf{x}).$$

Information about the rate of change of $\mathbf{y}$ with respect to $\mathbf{x}$ is contained in the various partial derivatives $\partial y_i/\partial x_j$, $(1 \le i \le m,\ 1 \le j \le n)$, and is conveniently organized into an $m \times n$ matrix, $D\mathbf{f}(\mathbf{x})$, called the **Jacobian matrix** of the transformation $\mathbf{f}$:

$$D\mathbf{f}(\mathbf{x}) = \begin{pmatrix} \dfrac{\partial y_1}{\partial x_1} & \dfrac{\partial y_1}{\partial x_2} & \cdots & \dfrac{\partial y_1}{\partial x_n} \\ \dfrac{\partial y_2}{\partial x_1} & \dfrac{\partial y_2}{\partial x_2} & \cdots & \dfrac{\partial y_2}{\partial x_n} \\ \vdots & \vdots & & \vdots \\ \dfrac{\partial y_m}{\partial x_1} & \dfrac{\partial y_m}{\partial x_2} & \cdots & \dfrac{\partial y_m}{\partial x_n} \end{pmatrix}$$

The linear transformation (see Section 10.6) represented by the Jacobian matrix is called **the derivative** of the transformation $\mathbf{f}$.

Remark We can regard the scalar-valued function of two variables, $f(x, y)$ say, as a transformation from $\mathbb{R}^2$ to $\mathbb{R}$. Its derivative is then the linear transformation with matrix

$$Df(x, y) = \big(f_1(x, y),\ f_2(x, y)\big).$$

It is not our purpose to enter into a study of such *vector-valued functions of a vector variable* at this point, but we can observe here that the Jacobian matrix of the composition of two such transformations is the matrix product of their Jacobian matrices.

To see this, let $\mathbf{y} = \mathbf{f}(\mathbf{x})$ be a transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ as described above, and let $\mathbf{z} = \mathbf{g}(\mathbf{y})$ be another such transformation from $\mathbb{R}^m$ to $\mathbb{R}^k$ given by

$$\begin{aligned} z_1 &= g_1(y_1, y_2, \ldots, y_m) \\ z_2 &= g_2(y_1, y_2, \ldots, y_m) \\ &\vdots \\ z_k &= g_k(y_1, y_2, \ldots, y_m), \end{aligned}$$

which has the $k \times m$ Jacobian matrix

$$D\mathbf{g}(\mathbf{y}) = \begin{pmatrix} \frac{\partial z_1}{\partial y_1} & \frac{\partial z_1}{\partial y_2} & \cdots & \frac{\partial z_1}{\partial y_m} \\ \frac{\partial z_2}{\partial y_1} & \frac{\partial z_2}{\partial y_2} & \cdots & \frac{\partial z_2}{\partial y_m} \\ \vdots & \vdots & & \vdots \\ \frac{\partial z_k}{\partial y_1} & \frac{\partial z_k}{\partial y_2} & \cdots & \frac{\partial z_k}{\partial y_m} \end{pmatrix}.$$

Then the composition $\mathbf{z} = \mathbf{g} \circ \mathbf{f}(\mathbf{x}) = \mathbf{g}(\mathbf{f}(\mathbf{x}))$ given by

$$\begin{aligned} z_1 &= g_1(f_1(x_1, \ldots, x_n), \ldots, f_m(x_1, \ldots, x_n)) \\ z_2 &= g_2(f_1(x_1, \ldots, x_n), \ldots, f_m(x_1, \ldots, x_n)) \\ &\vdots \\ z_k &= g_k(f_1(x_1, \ldots, x_n), \ldots, f_m(x_1, \ldots, x_n)) \end{aligned}$$

has, according to the Chain Rule, the $k \times n$ Jacobian matrix

$$\begin{pmatrix} \frac{\partial z_1}{\partial x_1} & \frac{\partial z_1}{\partial x_2} & \cdots & \frac{\partial z_1}{\partial x_n} \\ \frac{\partial z_2}{\partial x_1} & \frac{\partial z_2}{\partial x_2} & \cdots & \frac{\partial z_2}{\partial x_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial z_k}{\partial x_1} & \frac{\partial z_k}{\partial x_2} & \cdots & \frac{\partial z_k}{\partial x_n} \end{pmatrix} = \begin{pmatrix} \frac{\partial z_1}{\partial y_1} & \frac{\partial z_1}{\partial y_2} & \cdots & \frac{\partial z_1}{\partial y_m} \\ \frac{\partial z_2}{\partial y_1} & \frac{\partial z_2}{\partial y_2} & \cdots & \frac{\partial z_2}{\partial y_m} \\ \vdots & \vdots & & \vdots \\ \frac{\partial z_k}{\partial y_1} & \frac{\partial z_k}{\partial y_2} & \cdots & \frac{\partial z_k}{\partial y_m} \end{pmatrix} \begin{pmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} & \cdots & \frac{\partial y_1}{\partial x_n} \\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} & \cdots & \frac{\partial y_2}{\partial x_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial y_m}{\partial x_1} & \frac{\partial y_m}{\partial x_2} & \cdots & \frac{\partial y_m}{\partial x_n} \end{pmatrix}$$

This is, in fact, the Chain Rule for compositions of transformations:

$$D(\mathbf{g} \circ \mathbf{f})(\mathbf{x}) = D\mathbf{g}(\mathbf{f}(\mathbf{x}))D\mathbf{f}(\mathbf{x}),$$

and exactly mimics the one-variable Chain Rule $D(g \circ f)(x) = Dg(f(x))Df(x)$.

The transformation $\mathbf{y} = \mathbf{f}(\mathbf{x})$ also defines a vector $d\mathbf{y}$ of differentials of the variables y_i in terms of the vector $d\mathbf{x}$ of differentials of the variables x_j. Writing $d\mathbf{y}$ and $d\mathbf{x}$ as column vectors we have

$$d\mathbf{y} = \begin{pmatrix} dy_1 \\ dy_2 \\ \vdots \\ dy_m \end{pmatrix} = \begin{pmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} & \cdots & \frac{\partial y_1}{\partial x_n} \\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} & \cdots & \frac{\partial y_2}{\partial x_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial y_m}{\partial x_1} & \frac{\partial y_m}{\partial x_2} & \cdots & \frac{\partial y_m}{\partial x_n} \end{pmatrix} \begin{pmatrix} dx_1 \\ dx_2 \\ \vdots \\ dx_n \end{pmatrix} = D\mathbf{f}(\mathbf{x})d\mathbf{x}.$$

Example 4 Find the Jacobian matrix $D\mathbf{f}(1, 0)$ for the transformation from $\mathbb{R}^2$ to $\mathbb{R}^3$ given by

$$\mathbf{f}(x, y) = (xe^y + \cos(\pi y), x^2, x - e^y)$$

and use it to find an approximate value for $\mathbf{f}(1.02, 0.01)$.

Solution $D\mathbf{f}(x, y)$ is the 3×2 matrix whose jth row consists of the partial derivatives of the jth component of $\mathbf{f}$ with respect to x and y. Thus,

$$D\mathbf{f}(1,0) = \begin{pmatrix} e^y & xe^y - \pi\sin(\pi y) \\ 2x & 0 \\ 1 & -e^y \end{pmatrix}\Bigg|_{(1,0)} = \begin{pmatrix} 1 & 1 \\ 2 & 0 \\ 1 & -1 \end{pmatrix}.$$

Since $\mathbf{f}(1, 0) = (2, 1, 0)$ and $d\mathbf{x} = \begin{pmatrix} 0.02 \\ 0.01 \end{pmatrix}$, we have

$$d\mathbf{f} = D\mathbf{f}(1,0)\,d\mathbf{x} = \begin{pmatrix} 1 & 1 \\ 2 & 0 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 0.02 \\ 0.01 \end{pmatrix} = \begin{pmatrix} 0.03 \\ 0.04 \\ 0.01 \end{pmatrix}.$$

Therefore, $\mathbf{f}(1.02, 0.01) \approx (2.03, 1.04, 0.01)$. ∎

For transformations between spaces of the same dimension (say from $\mathbb{R}^n$ to $\mathbb{R}^n$), the corresponding Jacobian matrices are square and have determinants. These Jacobian determinants will play an important role in our consideration of implicit functions and inverse functions in Section 12.8 and in changes of variables in multiple integrals in Chapter 14.

Maple's **linalg** package has a function **jacobian** that takes two inputs, a vector of expressions and a vector of variables, and produces the Jacobian matrix of the partial derivatives of those expressions with respect to the variables. For example,

```
> with(linalg):
> jacobian([x*y*exp(z), (x+2*y)*cos(z)],[x,y,z]);
```

$$\begin{bmatrix} ye^z & xe^z & xye^z \\ \cos(z) & 2\cos(z) & -(x+2y)\sin(z) \end{bmatrix}$$

Exercises 12.6

In Exercises 1–6, use suitable linearizations to find approximate values for the given functions at the points indicated.

1. $f(x, y) = x^2y^3$ at $(3.1, 0.9)$
2. $f(x, y) = \tan^{-1}\left(\dfrac{y}{x}\right)$ at $(3.01, 2.99)$
3. $f(x, y) = \sin(\pi xy + \ln y)$ at $(0.01, 1.05)$
4. $f(x, y) = \dfrac{24}{x^2 + xy + y^2}$ at $(2.1, 1.8)$
5. $f(x, y, z) = \sqrt{x + 2y + 3z}$ at $(1.9, 1.8, 1.1)$
6. $f(x, y) = x\,e^{y+x^2}$ at $(2.05, -3.92)$
7. The edges of a rectangular box are each measured to within an accuracy of 1% of their values. What is the approximate maximum percentage error in
   (a) the calculated volume of the box,
   (b) the calculated area of one of the faces of the box, and
   (c) the calculated length of a diagonal of the box?
8. The radius and height of a right-circular conical tank are measured to be 25 ft and 21 ft, respectively. Each measurement is accurate to within 0.5 in. By about how much can the calculated volume of the tank be in error?
9. By approximately how much can the calculated area of the conical surface of the tank in Exercise 8 be in error?
10. Two sides and the contained angle of a triangular plot of land are measured to be 224 m, 158 m, and 64°, respectively. The length measurements were accurate to within 0.4 m and the angle measurement to within 2°. What is the approximate maximum percentage error if the area of the plot is calculated from these measurements?
11. The angle of elevation of the top of a tower is measured at two points A and B on the ground in the same direction from the base of the tower. The angles are 50° at A and 35° at B, each measured to within 1°. The distance AB is measured to be 100 m with error at most 0.1%. What is the calculated height of the building, and by about how much can it be in error? To which of the three measurements is the calculated height most sensitive?
12. By approximately what percentage will the value of $w = \dfrac{x^2y^3}{z^4}$ increase or decrease if x increases by 1%, y

increases by 2%, and z and increases by 3%?

13. Find the Jacobian matrix for the transformation $\mathbf{f}(r, \theta) = (x, y)$, where

$$x = r\cos\theta \quad \text{and} \quad y = r\sin\theta.$$

(Although (r, θ) can be regarded as *polar coordinates* in the xy-plane, they are Cartesian coordinates in their own $r\theta$-plane.)

14. Find the Jacobian matrix for the transformation $\mathbf{f}(\rho, \phi, \theta) = (x, y, z)$, where

$$x = \rho\sin\phi\cos\theta, \quad y = \rho\sin\phi\sin\theta, \quad z = \rho\cos\phi.$$

Here (ρ, ϕ, θ) are *spherical coordinates* in xyz-space. They will be formally introduced in Section 14.5.

15. Find the Jacobian matrix $D\mathbf{f}(x, y, z)$ for the transformation of $\mathbb{R}^3$ to $\mathbb{R}^2$ given by

$$\mathbf{f}(x, y, z) = (x^2 + yz, y^2 - x\ln z).$$

Use $D\mathbf{f}(2, 2, 1)$ to help you find an approximate value for $\mathbf{f}(1.98, 2.01, 1.03)$.

16. Find the Jacobian matrix $D\mathbf{g}(1, 3, 3)$ for the transformation of $\mathbb{R}^3$ to $\mathbb{R}^3$ given by

$$\mathbf{g}(r, s, t) = (r^2 s, r^2 t, s^2 - t^2)$$

and use the result to find an approximate value for $\mathbf{g}(0.99, 3.02, 2.97)$.

17. Prove that if $f(x, y)$ is differentiable at (a, b), then $f(x, y)$ is continuous at (a, b).

* **18.** Prove the following version of the Mean-Value Theorem: if $f(x, y)$ has first partial derivatives continuous near every point of the straight line segment joining the points (a, b) and $(a + h, b + k)$, then there exists a number θ satisfying $0 < \theta < 1$ such that

$$f(a+h, b+k) = f(a, b) + hf_1(a + \theta h, b + \theta k) + kf_2(a + \theta h, b + \theta k).$$

(*Hint:* apply the single-variable Mean-Value Theorem to $g(t) = f(a + th, b + tk)$.) Why could we not have used this result in place of Theorem 3 to prove Theorem 4 and hence the version of the Chain Rule given in this section?

* **19.** Generalize Exercise 18 as follows: show that, if $f(x, y)$ has continuous partial derivatives of second order near the point (a, b), then there exists a number θ satisfying $0 < \theta < 1$ such that, for h and k sufficiently small in absolute value,

$$\begin{aligned} f(a+h, b+k) = {} & f(a, b) + hf_1(a, b) + kf_2(a, b) \\ & + h^2 f_{11}(a + \theta h, b + \theta k) \\ & + 2hkf_{12}(a + \theta h, b + \theta k) \\ & + k^2 f_{22}(a + \theta h, b + \theta k). \end{aligned}$$

Hence, show that there is a constant K such that for all sufficiently small h and k,

$$\big|f(a+h, b+k) - f(a, b) - hf_1(a, b) - kf_2(a, b)\big| \le K(h^2 + k^2).$$

12.7 Gradients and Directional Derivatives

A first partial derivative of a function of several variables gives the rate of change of that function with respect to distance measured in the direction of one of the coordinate axes. In this section we will develop a method for finding the rate of change of such a function with respect to distance measured in *any direction* in the domain of the function.

To begin, it is useful to combine the first partial derivatives of a function into a single *vector function* called a **gradient**. For simplicity, we will develop and interpret the gradient for functions of two variables. Extension to functions of three or more variables is straightforward and will be discussed later in this section.

DEFINITION 6

At any point (x, y) where the first partial derivatives of the function $f(x, y)$ exist, we define the **gradient vector** $\nabla f(x, y) = \textbf{grad}\, f(x, y)$ by

$$\nabla f(x, y) = \textbf{grad}\, f(x, y) = f_1(x, y)\mathbf{i} + f_2(x, y)\mathbf{j}.$$

Recall that **i** and **j** denote the unit basis vectors from the origin to the points $(1, 0)$ and $(0, 1)$ respectively. The symbol ∇, called *del* or *nabla*, is a *vector differential operator*:

$$\nabla = \mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y}.$$

We can *apply* this operator to a function $f(x, y)$ by writing the operator to the left of the function. The result is the gradient of the function

$$\nabla f(x, y) = \left(\mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y}\right) f(x, y) = f_1(x, y)\mathbf{i} + f_2(x, y)\mathbf{j}.$$

We will make extensive use of the del operator in Chapter 16.

Example 1 If $f(x, y) = x^2 + y^2$, then $\nabla f(x, y) = 2x\mathbf{i} + 2y\mathbf{j}$. In particular, $\nabla f(1, 2) = 2\mathbf{i} + 4\mathbf{j}$. Observe that this vector is perpendicular to the tangent line $x + 2y = 5$ to the circle $x^2 + y^2 = 5$ at $(1, 2)$. This circle is the level curve of f that passes through the point $(1, 2)$. (See Figure 12.23.) As the following theorem shows, this perpendicularity is not a coincidence. ■

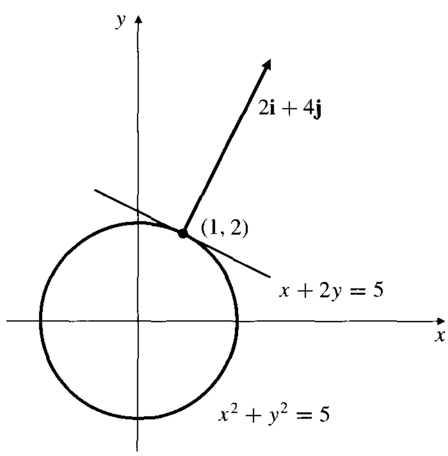

Figure 12.23 The gradient of $f(x, y) = x^2 + y^2$ at $(1, 2)$ is normal to the level curve of f through $(1, 2)$

THEOREM 6

If $f(x, y)$ is differentiable at the point (a, b) and $\nabla f(a, b) \neq \mathbf{0}$, then $\nabla f(a, b)$ is a normal vector to the level curve of f that passes through (a, b).

PROOF Let $\mathbf{r} = \mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j}$ be a parametrization of the level curve of f such that $x(0) = a$ and $y(0) = b$. Then for all t near 0, $f\big(x(t), y(t)\big) = f(a, b)$. Differentiating this equation with respect to t using the Chain Rule, we obtain

$$f_1\big(x(t), y(t)\big)\frac{dx}{dt} + f_2\big(x(t), y(t)\big)\frac{dy}{dt} = 0.$$

At $t = 0$ this says that $\nabla f(a, b) \bullet \left.\dfrac{d\mathbf{r}}{dt}\right|_{t=0} = 0$, that is, ∇f is perpendicular to the tangent vector $d\mathbf{r}/dt$ to the level curve at (a, b). ●

Directional Derivatives

The first partial derivatives $f_1(a, b)$ and $f_2(a, b)$ give the rates of change of $f(x, y)$ at (a, b) measured in the directions of the positive x- and y-axes, respectively. If we want to know how fast $f(x, y)$ changes value as we move through the domain of f at (a, b) in some other direction, we require a more general **directional derivative**. We can specify the direction by means of a nonzero vector. It is most convenient to use a *unit vector*.

DEFINITION 7

Let $\mathbf{u} = u\mathbf{i} + v\mathbf{j}$ be a unit vector, so that $u^2 + v^2 = 1$. The **directional derivative** of $f(x, y)$ at (a, b) in the direction of $\mathbf{u}$ is the rate of change of $f(x, y)$ with respect to distance measured at (a, b) along a ray in the direction of $\mathbf{u}$ in the xy-plane. (See Figure 12.24.) This directional derivative is given by

$$D_{\mathbf{u}}f(a, b) = \lim_{h\to 0+} \frac{f(a + hu, b + hv) - f(a, b)}{h}.$$

It is also given by

$$D_{\mathbf{u}}f(a, b) = \frac{d}{dt} f(a + tu, b + tv)\Big|_{t=0}$$

if the derivative on the right side exists.

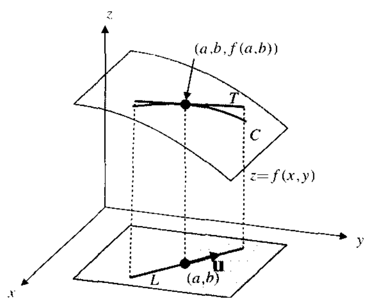

Figure 12.24 Unit vector $\mathbf{u}$ determines a line L through (a, b) in the domain of f. The vertical plane containing L intersects the graph of f in a curve C whose tangent T at $(a, b, f(a, b))$ has slope $D_{\mathbf{u}}f(a, b)$

Observe that directional derivatives in directions parallel to the coordinate axes are given directly by first partials: $D_{\mathbf{i}}f(a, b) = f_1(a, b)$, $D_{\mathbf{j}}f(a, b) = f_2(a, b)$, $D_{-\mathbf{i}}f(a, b) = -f_1(a, b)$, and $D_{-\mathbf{j}}f(a, b) = -f_2(a, b)$. The following theorem shows how the gradient can be used to calculate any directional derivative.

THEOREM 7

Using the gradient to find directional derivatives

If f is differentiable at (a, b) and $\mathbf{u} = u\mathbf{i} + v\mathbf{j}$ is a unit vector, then the directional derivative of f at (a, b) in the direction of $\mathbf{u}$ is given by

$$D_{\mathbf{u}}f(a, b) = \mathbf{u} \bullet \nabla f(a, b).$$

PROOF By the Chain Rule:

$$\begin{aligned} D_{\mathbf{u}}f(a, b) &= \frac{d}{dt} f(a + tu, b + tv)\Big|_{t=0} \\ &= uf_1(a, b) + vf_2(a, b) = \mathbf{u} \bullet \nabla f(a, b). \end{aligned}$$

We already know that having partial derivatives at a point does not imply that a function is continuous there, let alone that it is differentiable. The same can be said about directional derivatives. It is possible for a function to have a directional derivative in every direction at a given point and still not be continuous at that point. See Exercise 37 below for an example of such a function.

Given any nonzero vector $\mathbf{v}$, we can always obtain a unit vector in the same direction by dividing $\mathbf{v}$ by its length. The directional derivative of f at (a, b) in the direction of $\mathbf{v}$ is therefore given by

$$D_{\mathbf{v}/|\mathbf{v}|} f(a, b) = \frac{\mathbf{v}}{|\mathbf{v}|} \bullet \nabla f(a, b).$$

Example 2 Find the rate of change of $f(x, y) = y^4 + 2xy^3 + x^2y^2$ at $(0, 1)$ measured in each of the following directions:

(a) $\mathbf{i} + 2\mathbf{j}$, (b) $\mathbf{j} - 2\mathbf{i}$, (c) $3\mathbf{i}$, (d) $\mathbf{i} + \mathbf{j}$.

Solution We calculate

$$\nabla f(x, y) = (2y^3 + 2xy^2)\mathbf{i} + (4y^3 + 6xy^2 + 2x^2y)\mathbf{j},$$
$$\nabla f(0, 1) = 2\mathbf{i} + 4\mathbf{j}.$$

(a) The directional derivative of f at $(0, 1)$ in the direction of $\mathbf{i} + 2\mathbf{j}$ is

$$\frac{\mathbf{i} + 2\mathbf{j}}{|\mathbf{i} + 2\mathbf{j}|} \bullet (2\mathbf{i} + 4\mathbf{j}) = \frac{2 + 8}{\sqrt{5}} = 2\sqrt{5}.$$

Observe that $\mathbf{i} + 2\mathbf{j}$ points in the same direction as $\nabla f(0, 1)$ so the directional derivative is positive and equal to the length of $\nabla f(0, 1)$.

(b) The directional derivative of f at $(0, 1)$ in the direction of $\mathbf{j} - 2\mathbf{i}$ is

$$\frac{-2\mathbf{i} + \mathbf{j}}{|-2\mathbf{i} + \mathbf{j}|} \bullet (2\mathbf{i} + 4\mathbf{j}) = \frac{-4 + 4}{\sqrt{5}} = 0.$$

Since $\mathbf{j} - 2\mathbf{i}$ is perpendicular to $\nabla f(0, 1)$, it is tangent to the level curve of f through $(0, 1)$, so the directional derivative in that direction is zero.

(c) The directional derivative of f at $(0, 1)$ in the direction of $3\mathbf{i}$ is

$$\mathbf{i} \bullet (2\mathbf{i} + 4\mathbf{j}) = 2.$$

As noted previously, the directional derivative of f in the direction of the positive x-axis is just $f_1(0, 1)$.

(d) The directional derivative of f at $(0, 1)$ in the direction of $\mathbf{i} + \mathbf{j}$ is

$$\frac{\mathbf{i} + \mathbf{j}}{|\mathbf{i} + \mathbf{j}|} \bullet (2\mathbf{i} + 4\mathbf{j}) = \frac{2 + 4}{\sqrt{2}} = 3\sqrt{2}.$$

If we move along the surface $z = f(x, y)$ through the point $(0, 1, 1)$ in a direction making horizontal angles of 45° with the positive directions of the x- and y-axes, we would be rising at a rate of $3\sqrt{2}$ vertical units per horizontal unit moved. ∎

Remark A direction in the plane can be specified by a polar angle. The direction making angle ϕ with the positive direction of the x-axis corresponds to the unit vector (see Figure 12.25)

$$\mathbf{u}_\phi = \cos\phi\mathbf{i} + \sin\phi\mathbf{j},$$

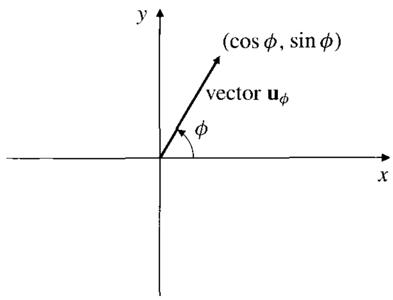

Figure 12.25 The unit vector specified by a polar angle ϕ

so the directional derivative of f at (x, y) in that direction is

$$D_\phi f(x, y) = D_{\mathbf{u}_\phi} f(x, y) = \mathbf{u}_\phi \bullet \nabla f(x, y) = f_1(x, y)\cos\phi + f_2(x, y)\sin\phi.$$

Note the use of the symbol $D_\phi f(x, y)$ to denote a derivative of f with respect to *distance* measured in the direction ϕ.

As observed in the previous example, Theorem 7 provides a useful interpretation for the gradient vector. For any unit vector $\mathbf{u}$ we have

$$D_{\mathbf{u}} f(a, b) = \mathbf{u} \bullet \nabla f(a, b) = |\nabla f(a, b)| \cos\theta,$$

where θ is the angle between the vectors $\mathbf{u}$ and $\nabla f(a, b)$. Since $\cos\theta$ only takes on values between -1 and 1, $D_{\mathbf{u}} f(a, b)$ only takes on values between $-|\nabla f(a, b)|$ and $|\nabla f(a, b)|$. Moreover, $D_{\mathbf{u}} f(a, b) = -|\nabla f(a, b)|$ if and only if $\mathbf{u}$ points in the opposite direction to $\nabla f(a, b)$ (so that $\cos\theta = -1$), and $D_{\mathbf{u}} f(a, b) = |\nabla f(a, b)|$ if and only if $\mathbf{u}$ points in the same direction as $\nabla f(a, b)$ (so that $\cos\theta = 1$). The directional derivative is zero in the direction $\theta = \pi/2$; this is the direction of the (tangent line to the) level curve of f through (a, b).

We summarize these properties of the gradient as follows:

Geometric properties of the gradient vector

(i) At (a, b), $f(x, y)$ increases most rapidly in the direction of the gradient vector $\nabla f(a, b)$. The maximum rate of increase is $|\nabla f(a, b)|$.

(ii) At (a, b), $f(x, y)$ decreases most rapidly in the direction of $-\nabla f(a, b)$. The maximum rate of decrease is $|\nabla f(a, b)|$.

(iii) The rate of change of $f(x, y)$ at (a, b) is zero in directions tangent to the level curve of f that passes through (a, b).

Look again at the topographic map in Figure 12.6 in Section 12.1. The streams on the map flow in the direction of steepest descent, that is, in the direction of $-\nabla f$, where f measures the elevation of land. The streams therefore cross the contours (the level curves of f) at right angles. Like the stream, an experienced skier might choose a downhill path close to the direction of the negative gradient, while a novice skier would prefer to stay closer to the level curves.

Example 3 The temperature at position (x, y) in a region of the xy-plane is T°C, where

$$T(x, y) = x^2 e^{-y}.$$

In what direction at the point $(2, 1)$ does the temperature increase most rapidly? What is the rate of increase of f in that direction?

Solution We have

$$\nabla T(x, y) = 2x\, e^{-y}\mathbf{i} - x^2 e^{-y}\mathbf{j},$$
$$\nabla T(2, 1) = \frac{4}{e}\mathbf{i} - \frac{4}{e}\mathbf{j} = \frac{4}{e}(\mathbf{i} - \mathbf{j}).$$

At $(2, 1)$, $T(x, y)$ increases most rapidly in the direction of the vector $\mathbf{i} - \mathbf{j}$. The rate of increase in this direction is $|\nabla T(2, 1)| = 4\sqrt{2}/e$ °C/unit distance. ■

Example 4 A hiker is standing beside a stream on the side of a mountain, examining her map of the region. The height of land (in metres) at any point (x, y) is given by

$$h(x, y) = \frac{20{,}000}{3 + x^2 + 2y^2},$$

where x and y (in kilometres) denote the coordinates of the point on the hiker's map. The hiker is at the point $(3, 2)$.

(a) What is the direction of flow of the stream at $(3, 2)$ on the hiker's map? How fast is the stream descending at her location?
and showing the stream.

Solution

(a) We begin by calculating the gradient of h and its length at $(3, 2)$:

$$\nabla h(x, y) = -\frac{20{,}000}{(3 + x^2 + 2y^2)^2}(2x\mathbf{i} + 4y\mathbf{j}),$$
$$\nabla h(3, 2) = -100(3\mathbf{i} + 4\mathbf{j}),$$
$$|\nabla h(3, 2)| = 500.$$

The stream is flowing in the direction whose horizontal projection at $(3, 2)$ is $-\nabla h(3, 2)$, that is, in the horizontal direction of the vector $3\mathbf{i} + 4\mathbf{j}$. The stream is descending at a rate of 500 m/km, that is, 0.5 metres per horizontal metre travelled.

(b) Coordinates on the map are the coordinates (x, y) in the domain of the height function h. We can find an equation of the path of the stream on a map of the region by setting up a differential equation for a change of position along the path. If the vector $d\mathbf{r} = dx\,\mathbf{i} + dy\,\mathbf{j}$ is tangent to the path of the stream at point (x, y) on the map, then $d\mathbf{r}$ is parallel to $\nabla h(x, y)$. Hence, the components of these two vectors are proportional:

$$\frac{dx}{2x} = \frac{dy}{4y} \quad \text{or} \quad \frac{dy}{y} = \frac{2dx}{x}.$$

Integrating both sides of this equation, we get $\ln y = 2 \ln x + \ln C$, or $y = Cx^2$. Since the path of the stream passes through $(3, 2)$, we have $C = 2/9$ and the equation is $9y = 2x^2$.

(c) Suppose the hiker moves away from $(3, 2)$ in the direction of the unit vector $\mathbf{u}$. She will be ascending at an inclination of $15°$ if the directional derivative of h in the direction of $\mathbf{u}$ is $1{,}000 \tan 15° \approx 268$. (The 1,000 compensates for the fact that the vertical units are metres while the horizontal units are kilometres.) If θ is the angle between $\mathbf{u}$ and the upstream direction, then

$$500 \cos\theta = |\nabla h(3, 2)| \cos\theta = D_{\mathbf{u}}h(3, 2) \approx 268.$$

Hence $\cos\theta \approx 0.536$ and $\theta \approx 57.6°$. She should set out in a direction making a horizontal angle of about $58°$ with the upstream direction.

(d) A suitable sketch of the map is given in Figure 12.26. ■

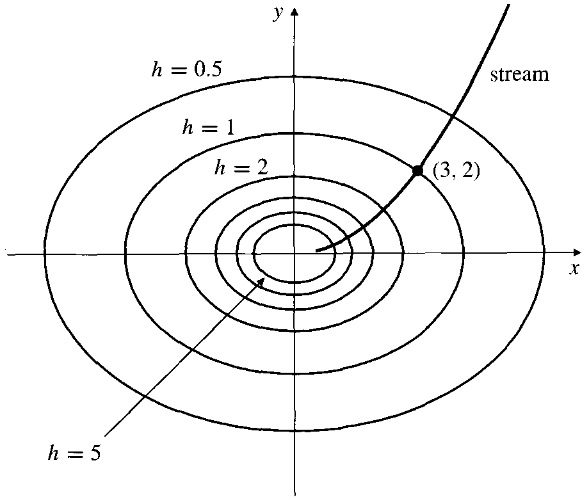

Figure 12.26 The hiker's map. Unlike most mountains, this one has perfectly elliptical contours

Example 5 Find the second directional derivative of $f(x, y)$ in the direction making angle ϕ with the positive x-axis.

Solution As observed earlier, the first directional derivative is

$$D_\phi f(x, y) = (\cos\phi\mathbf{i} + \sin\phi\mathbf{j}) \bullet \nabla f(x, y) = f_1(x, y)\cos\phi + f_2(x, y)\sin\phi.$$

The second directional derivative is therefore

$$\begin{aligned} D_\phi^2 f(x, y) &= D_\phi\big(D_\phi f(x, y)\big) \\ &= (\cos\phi\mathbf{i} + \sin\phi\mathbf{j}) \bullet \nabla\Big(f_1(x, y)\cos\phi + f_2(x, y)\sin\phi\Big) \\ &= \Big(f_{11}(x, y)\cos\phi + f_{21}(x, y)\sin\phi\Big)\cos\phi \\ &\quad + \Big(f_{12}(x, y)\cos\phi + f_{22}(x, y)\sin\phi\Big)\sin\phi \\ &= f_{11}(x, y)\cos^2\phi + 2f_{12}(x, y)\cos\phi\sin\phi + f_{22}(x, y)\sin^2\phi. \end{aligned}$$

Note that if $\phi = 0$ or $\phi = \pi$ (so the directional derivative is in a direction parallel to the x-axis) then $D_\phi^2 f(x, y) = f_{11}(x, y)$. Similarly, $D_\phi^2 f(x, y) = f_{22}(x, y)$ if $\phi = \pi/2$ or $3\pi/2$.

Rates Perceived by a Moving Observer

Suppose that an observer is moving around in the xy-plane measuring the value of a function $f(x, y)$ defined in the plane as he passes through each point (x, y). (For instance, $f(x, y)$ might be the temperature at (x, y).) If the observer is moving with velocity $\mathbf{v}$ at the instant when he passes through the point (a, b), how fast would he observe $f(x, y)$ to be changing at that moment?

At the moment in question the observer is moving in the direction of the unit vector $\mathbf{v}/|\mathbf{v}|$. The rate of change of $f(x, y)$ at (a, b) in that direction is

$$D_{\mathbf{v}/|\mathbf{v}|} f(a, b) = \frac{\mathbf{v}}{|\mathbf{v}|} \bullet \nabla f(a, b)$$

measured in units of f per unit distance in the xy-plane. To convert this rate to units of f per unit time, we must multiply by the speed of the observer, $|\mathbf{v}|$ units of distance per unit time. Thus, the time rate of change of $f(x, y)$ as measured by the observer passing through (a, b) is

$$|\mathbf{v}| \frac{\mathbf{v}}{|\mathbf{v}|} \bullet \nabla f(a, b) = \mathbf{v} \bullet \nabla f(a, b).$$

It is natural to extend our use of the symbol $D_{\mathbf{v}} f(a, b)$ to represent this rate even though $\mathbf{v}$ is not (necessarily) a unit vector. Thus

The rate of change of $f(x, y)$ at (a, b) as measured by an observer moving through (a, b) with velocity $\mathbf{v}$ is

$$D_{\mathbf{v}} f(a, b) = \mathbf{v} \bullet \nabla f(a, b)$$

units of f per unit time.

If the hiker in Example 4 moves away from (3, 2) with horizontal velocity $\mathbf{v} = -\mathbf{i} - \mathbf{j}$ km/h, then she will be rising at a rate of

$$\mathbf{v} \bullet \nabla h(3, 2) = (-\mathbf{i} - \mathbf{j}) \bullet \left(-\frac{1}{10}(3\mathbf{i} + 4\mathbf{j}) \right) = \frac{7}{10} \text{ km/h}.$$

As defined here, $D_{\mathbf{v}} f$ is the spatial component of the derivative of f following the motion. See Example 6 in Section 12.5. The rate of change of the reading on the moving thermometer in that example can be expressed as

$$\frac{dT}{dt} = D_{\mathbf{v}} T(x, y, z, t) + \frac{\partial T}{\partial t},$$

where $\mathbf{v}$ is the velocity of the moving thermometer and $D_{\mathbf{v}} T = \mathbf{v} \bullet \nabla T$. The gradient is being taken with respect to the *three spatial variables* only. (See below for the gradient in 3-space.)

The Gradient in Three and More Dimensions

By analogy with the two-dimensional case, a function $f(x_1, x_2, \ldots, x_n)$ of n variables possessing first partial derivatives has gradient given by

$$\nabla f(x_1, x_2, \ldots, x_n) = \frac{\partial f}{\partial x_1} \mathbf{e_1} + \frac{\partial f}{\partial x_2} \mathbf{e_2} + \cdots + \frac{\partial f}{\partial x_n} \mathbf{e_n},$$

where $\mathbf{e}_j$ is the unit vector from the origin to the unit point on the jth coordinate axis. In particular, for a function of three variables,

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

The level surface of $f(x, y, z)$ passing through (a, b, c) has a tangent plane there if f is differentiable at (a, b, c) and $\nabla f(a, b, c) \neq \mathbf{0}$.

For functions of any number of variables, the vector $\nabla f(P_0)$ is normal to the "level surface" of f passing through the point P_0 (i.e., the (hyper)surface with equation $f(x_1, \ldots, x_n) = f(P_0)$), and, if f is differentiable at P_0, the rate of change of f at P_0 in the direction of the unit vector $\mathbf{u}$ is given by $\mathbf{u} \bullet \nabla f(P_0)$. Equations of tangent planes to surfaces in 3-space can be found easily with the aid of gradients.

Example 6 Let $f(x, y, z) = x^2 + y^2 + z^2$.

(a) Find $\nabla f(x, y, z)$ and $\nabla f(1, -1, 2)$.

(b) Find an equation of the tangent plane to the sphere $x^2 + y^2 + z^2 = 6$ at the point $(1, -1, 2)$.

(c) What is the maximum rate of increase of f at $(1, -1, 2)$?

(d) What is the rate of change with respect to distance of f at $(1, -1, 2)$ measured in the direction from that point toward the point $(3, 1, 1)$?

Solution

(a) $\nabla f(x, y, z) = 2x\mathbf{i} + 2y\mathbf{j} + 2z\mathbf{k}$, so $\nabla f(1, -1, 2) = 2\mathbf{i} - 2\mathbf{j} + 4\mathbf{k}$.

(b) The required tangent plane has $\nabla f(1, -1, 2)$ as normal. (See Figure 12.27(a).) Therefore, its equation is given by $2(x-1) - 2(y+1) + 4(z-2) = 0$ or, more simply, $x - y + 2z = 6$.

(c) The maximum rate of increase of f at $(1, -1, 2)$ is $|\nabla f(1, -1, 2)| = 2\sqrt{6}$, and it occurs in the direction of the vector $\mathbf{i} - \mathbf{j} + 2\mathbf{k}$.

(d) The direction from $(1, -1, 2)$ toward $(3, 1, 1)$ is specified by $2\mathbf{i} + 2\mathbf{j} - \mathbf{k}$. The rate of change of f with respect to distance in this direction is

$$\frac{2\mathbf{i} + 2\mathbf{j} - \mathbf{k}}{\sqrt{4+4+1}} \bullet (2\mathbf{i} - 2\mathbf{j} + 4\mathbf{k}) = \frac{4-4-4}{3} = -\frac{4}{3},$$

that is, f decreases at rate $4/3$ of a unit per horizontal unit moved. ■

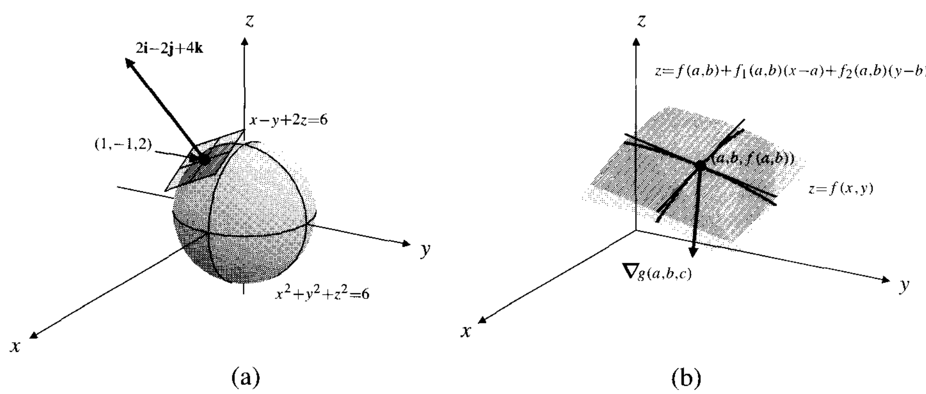

Figure 12.27

(a) The tangent plane to $x^2 + y^2 + z^2 = 6$ at $(1, -1, 2)$

(b) The gradient of $f(x, y) - z$ at $(a, b, f(a, b))$ is normal to the tangent plane to $z = f(x, y)$ at that point

BEWARE

Make sure you understand the difference between the graph of a function and a level curve or level surface of that function. (See the discussion following this example.) Here, the surface $z = f(x, y)$ is the *graph* of the function f, but it is also a *level surface* of a *different* function g.

Example 7 The graph of a *function* $f(x, y)$ of two variables is the graph of the *equation* $z = f(x, y)$ in 3-space. This surface is also the level surface $g(x, y, z) = 0$ of the 3-variable function

$$g(x, y, z) = f(x, y) - z.$$

If f is differentiable at (a, b) and $c = f(a, b)$, then g is differentiable at (a, b, c), and

$$\nabla g(a, b, c) = f_1(a, b)\mathbf{i} + f_2(a, b)\mathbf{j} - \mathbf{k}$$

is normal to $g(x, y, z) = 0$ at (a, b, c). (Note that $\nabla g(a, b, c) \neq \mathbf{0}$, since its z component is -1.) It follows that the graph of f has nonvertical tangent plane at (a, b) given by

$$f_1(a, b)(x - a) + f_2(a, b)(y - b) - (z - c) = 0,$$

or

$$z = f(a, b) + f_1(a, b)(x - a) + f_2(a, b)(y - b).$$

(See Figure 12.27(b).) This result was obtained by a different argument in Section 12.3.

Students sometimes confuse graphs of functions with level curves or surfaces of those functions. In the above example we are talking about a *level surface* of the function $g(x, y, z)$ that happens to coincide with the *graph* of a different function, $f(x, y)$. Do not confuse that surface with the graph of g, which is a three-dimensional *hypersurface* in 4-space having equation $w = g(x, y, z)$. Similarly, do not confuse the tangent *plane* to the graph of $f(x, y)$ (i.e., the plane obtained in the above example) with the tangent *line* to the level curve of $f(x, y)$ passing through (a, b) and lying in the xy-plane. This line has an equation involving only x and y, namely, $f_1(a, b)(x - a) + f_2(a, b)(y - b) = 0$.

Example 8 Find a vector tangent to the curve of intersection of the two surfaces

$$z = x^2 - y^2 \qquad \text{and} \qquad xyz + 30 = 0$$

at the point $(-3, 2, 5)$.

Solution The coordinates of the given point satisfy the equations of both surfaces so the point lies on the curve of intersection of the two surfaces. A vector tangent to this curve at that point will be perpendicular to the normals to both surfaces, that is, to the vectors

$$\mathbf{n}_1 = \nabla(x^2 - y^2 - z)\Big|_{(-3,2,5)} = 2x\mathbf{i} - 2y\mathbf{j} - \mathbf{k}\Big|_{(-3,2,5)} = -6\mathbf{i} - 4\mathbf{j} - \mathbf{k},$$

$$\mathbf{n}_2 = \nabla(xyz + 30)\Big|_{(-3,2,5)} = (yz\mathbf{i} + xz\mathbf{j} + xy\mathbf{k})\Big|_{(-3,2,5)} = 10\mathbf{i} - 15\mathbf{j} - 6\mathbf{k}.$$

For the tangent vector $\mathbf{T}$ we can therefore use the cross product of these normals:

$$\mathbf{T} = \mathbf{n}_1 \times \mathbf{n}_2 = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ -6 & -4 & -1 \\ 10 & -15 & -6 \end{vmatrix} = 9\mathbf{i} - 46\mathbf{j} + 130\mathbf{k}.$$

Remark Maple's **linalg** package defines a function **grad** that takes a pair of arguments, an expression and a vector of variables, and produces the gradient vector of the expression with respect to those variables:

```
> f := (x,y,z) -> x*exp(y)/z;
```

$$f := (x, y, z) \to \frac{x e^y}{z}$$

```
> with(linalg): G := grad(f(x,y,z), [x,y,z]);
```

$$G := \left[\frac{e^y}{z}, \frac{x e^y}{z}, -\frac{x e^y}{z^2}\right]$$

Getting the value of this gradient at a particular point, say (2, 0, 1), is surprisingly complicated. We must use the Maple evaluation function, **eval**, on G, then use the substitution function, **subs**, to substitute in the values for x, y, and z, and finally **simplify** the result.

```
> simplify(subs(x=2, y=0, z=1, eval(G)));
```

$$[1, 2, -2]$$

There is an easier way to accomplish this. We can define a gradient *operator*, let us call it Grad3, so that it acts on a function of three variables to produce a 3-vector-valued function.

```
> Grad3 := u -> [D[1](u), D[2](u), D[3](u)];
```

$$Grad3 := u \to [D_1(u), D_2(u), D_3(u)]$$

Note that Grad3 takes a *function* as its argument, not an expression.

```
> Grad3(f)(2,0,1);
```

$$[1, 2, -2]$$

To force evaluation of the result in decimal form, you can use **evalf** or just put decimal points in the coordinates:

```
> Grad3(f)(1,1,1); Grad3(f)(1.0,1.0,1.0);
```

$$[e, e, -e]$$

$$[2.718281828, 2.718281828, -2.718281828]$$

The definition of the gradient operator, Grad3, makes no use of the Maple linalg package and can be extended to apply to functions of different numbers of variables. We will use this approach again in Chapter 16 to define divergence and curl operators.

Exercises 12.7

In Exercises 1–6, find:

(a) the gradient of the given function at the point indicated,

(b) an equation of the plane tangent to the graph of the given function at the point whose x and y coordinates are given, and

(c) an equation of the straight line tangent, at the given point, to the level curve of the given function passing through that point.

1. $f(x, y) = x^2 - y^2$ at $(2, -1)$

2. $f(x, y) = \dfrac{x - y}{x + y}$ at $(1, 1)$

3. $f(x, y) = \dfrac{x}{x^2 + y^2}$ at $(1, 2)$

4. $f(x, y) = e^{xy}$ at $(2, 0)$

5. $f(x, y) = \ln(x^2 + y^2)$ at $(1, -2)$

6. $f(x, y) = \sqrt{1 + xy^2}$ at $(2, -2)$

In Exercises 7–9, find an equation of the tangent plane to the level surface of the given function that passes through the given point.

7. $f(x, y, z) = x^2 y + y^2 z + z^2 x$ at $(1, -1, 1)$

8. $f(x, y, z) = \cos(x + 2y + 3z)$ at $\left(\frac{\pi}{2}, \pi, \pi\right)$

9. $f(x, y, z) = y e^{-x^2} \sin z$ at $(0, 1, \pi/3)$

In Exercises 10–13, find the rate of change of the given function at the given point in the specified direction.

10. $f(x, y) = 3x - 4y$ at $(0, 2)$ in the direction of the vector $-2\mathbf{i}$

11. $f(x, y) = x^2 y$ at $(-1, -1)$ in the direction of the vector $\mathbf{i} + 2\mathbf{j}$

12. $f(x, y) = \dfrac{x}{1+y}$ at $(0, 0)$ in the direction of the vector $\mathbf{i} - \mathbf{j}$

13. $f(x, y) = x^2 + y^2$ at $(1, -2)$ in the direction making a (positive) angle of 60° with the positive x-axis

14. Let $f(x, y) = \ln|\mathbf{r}|$ where $\mathbf{r} = x\mathbf{i} + y\mathbf{j}$. Show that $\nabla f = \dfrac{\mathbf{r}}{|\mathbf{r}|^2}$.

15. Let $f(x, y, z) = |\mathbf{r}|^{-n}$, where $\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$. Show that $\nabla f = \dfrac{-n\mathbf{r}}{|\mathbf{r}|^{n+2}}$.

16. Show that, in terms of polar coordinates (r, θ) (where $x = r\cos\theta$ and $y = r\sin\theta$), the gradient of a function $f(r, \theta)$ is given by

$$\nabla f = \frac{\partial f}{\partial r}\hat{\mathbf{r}} + \frac{1}{r}\frac{\partial f}{\partial \theta}\hat{\boldsymbol{\theta}},$$

where $\hat{\mathbf{r}}$ is a unit vector in the direction of the position vector $\mathbf{r} = x\,\mathbf{i} + y\,\mathbf{j}$, and $\hat{\boldsymbol{\theta}}$ is a unit vector at right angles to $\hat{\mathbf{r}}$ in the direction of increasing θ.

17. In what directions at the point $(2, 0)$ does the function $f(x, y) = xy$ have rate of change -1? Are there directions in which the rate is -3? How about -2?

18. In what directions at the point (a, b, c) does the function $f(x, y, z) = x^2 + y^2 - z^2$ increase at half of its maximal rate at that point?

19. Find $\nabla f(a, b)$ for the differentiable function $f(x, y)$ given the directional derivatives

$$D_{(\mathbf{i}+\mathbf{j})/\sqrt{2}} f(a, b) = 3\sqrt{2} \text{ and } D_{(3\mathbf{i}-4\mathbf{j})/5} f(a, b) = 5.$$

20. If $f(x, y)$ is differentiable at (a, b), what condition should angles ϕ_1 and ϕ_2 satisfy in order that the gradient $\nabla f(a, b)$ can be determined from the values of the directional derivatives $D_{\phi_1} f(a, b)$ and $D_{\phi_2} f(a, b)$?

21. The temperature $T(x, y)$ at points of the xy-plane is given by $T(x, y) = x^2 - 2y^2$.

(a) Draw a contour diagram for T showing some isotherms (curves of constant temperature).

(b) In what direction should an ant at position $(2, -1)$ move if it wishes to cool off as quickly as possible?

(c) If the ant moves in that direction at speed k (units distance per unit time), at what rate does it experience the decrease of temperature?

(d) At what rate would the ant experience the decrease of temperature if it moved from $(2, -1)$ at speed k in the direction of the vector $-\mathbf{i} - 2\mathbf{j}$?

(e) Along what curve through $(2, -1)$ should the ant move in order to continue to experience maximum rate of cooling?

22. Find an equation of the curve in the xy-plane that passes through the point $(1, 1)$ and intersects all level curves of the function $f(x, y) = x^4 + y^2$ at right angles.

23. Find an equation of the curve in the xy-plane that passes through the point $(2, -1)$ and that intersects every curve with equation of the form $x^2y^3 = K$ at right angles.

24. Find the second directional derivative of $e^{-x^2-y^2}$ at the point $(a, b) \neq (0, 0)$ in the direction directly away from the origin.

25. Find the second directional derivative of $f(x, y, z) = xyz$ at $(2, 3, 1)$ in the direction of the vector $\mathbf{i} - \mathbf{j} - \mathbf{k}$.

26. Find a vector tangent to the curve of intersection of the two cylinders $x^2 + y^2 = 2$ and $y^2 + z^2 = 2$ at the point $(1, -1, 1)$.

27. Repeat Exercise 26 for the surfaces $x + y + z = 6$ and $x^2 + y^2 + z^2 = 14$ and the point $(1, 2, 3)$.

28. The temperature in 3-space is given by

$$T(x, y, z) = x^2 - y^2 + z^2 + xz^2.$$

At time $t = 0$ a fly passes through the point $(1, 1, 2)$, flying along the curve of intersection of the surfaces $z = 3x^2 - y^2$ and $2x^2 + 2y^2 - z^2 = 0$. If the fly's speed is 7, what rate of temperature change does it experience at $t = 0$?

29. State and prove a version of Theorem 6 for a function of three variables.

30. What is the level surface of $f(x, y, z) = \cos(x + 2y + 3z)$ that passes through (π, π, π)? What is the tangent plane to that level surface at that point? (Compare this exercise with Exercise 8 above.)

31. If $\nabla f(x, y) = 0$ throughout the disk $x^2 + y^2 < r^2$, prove that $f(x, y)$ is constant throughout the disk.

32. Theorem 6 implies that the level curve of $f(x, y)$ passing through (a, b) is smooth (has a tangent line) at (a, b) provided f is differentiable at (a, b) and satisfies $\nabla f(a, b) \neq \mathbf{0}$. Show that the level curve need not be smooth at (a, b) if $\nabla f(a, b) = \mathbf{0}$. (*Hint:* consider $f(x, y) = y^3 - x^2$ at $(0, 0)$.)

33. If $\mathbf{v}$ is a nonzero vector, express $D_{\mathbf{v}}(D_{\mathbf{v}} f)$ in terms of the components of $\mathbf{v}$ and the second partials of f. What is the interpretation of this quantity for a moving observer?

* **34.** An observer moves so that his position, velocity, and acceleration at time t are given by $\mathbf{r}(t) = x(t)\,\mathbf{i} + y(t)\,\mathbf{j} + z(t)\,\mathbf{k}$, $\mathbf{v}(t) = d\mathbf{r}/dt$, and $\mathbf{a}(t) = d\mathbf{v}/dt$. If the temperature in the vicinity of the observer depends only on position, $T = T(x, y, z)$, express the second time derivative of temperature as measured by the observer in terms of $D_{\mathbf{v}}$ and $D_{\mathbf{a}}$.

* **35.** Repeat Exercise 34 but with T depending explicitly on time as well as position: $T = T(x, y, z, t)$.

36. Let $f(x, y) = \begin{cases} \dfrac{\sin(xy)}{\sqrt{x^2+y^2}}, & \text{if } (x, y) \neq (0, 0) \\ 0, & \text{if } (x, y) = (0, 0). \end{cases}$

(a) Calculate $\nabla f(0, 0)$.

(b) Use the definition of directional derivative to calculate $D_{\mathbf{u}}f(0,0)$, where $\mathbf{u} = (\mathbf{i}+\mathbf{j})/\sqrt{2}$.

(b) Is $f(x, y)$ differentiable at $(0, 0)$? Why?

37. Let $f(x,y) = \begin{cases} 2x^2y/(x^4+y^2), & \text{if } (x,y) \neq (0,0) \\ 0, & \text{if } (x,y) = (0,0). \end{cases}$
Use the definition of directional derivative as a limit (Definition 7) to show that $D_{\mathbf{u}}f(0,0)$ exists for every unit vector $\mathbf{u} = u\mathbf{i} + v\mathbf{j}$ in the plane. Specifically, show that $D_{\mathbf{u}}f(0,0) = 0$ if $v = 0$, and $D_{\mathbf{u}}f(0,0) = 2u^2/v$ if $v \neq 0$. However, as was shown in Example 4 in Section 12.2, $f(x, y)$ has no limit as $(x, y) \to (0, 0)$, so it is not continuous there. Even if a function has directional derivatives in all directions at a point, it may not be continuous at that point.

38. Check that the Maple definition

```
> Grad := proc(u,n)
    options operator, arrow;
    local i;
    [seq(D[i](u),i=1..n)] end;
```

suitably defines a gradient operator for functions of n variables. Test it by defining a function $f(x, y, z, t) = xy^2z^3t^4$ and calculating `Grad(f,4)(1,1,1,1)`.

12.8 Implicit Functions

When we study the calculus of functions of one variable, we encounter examples of functions that are defined implicitly as solutions of equations in two variables. Suppose, for example, that $F(x, y) = 0$ is such an equation. Suppose that the point (a, b) satisfies the equation and that F has continuous first partial derivatives (and so is differentiable) at all points near (a, b). Can the equation be solved for y as a function of x near (a, b)? That is, does there exist a function $y(x)$ defined in some interval $I =]a-h, a+h[$ (where $h > 0$) satisfying $y(a) = b$ and such that

$$F\big(x, y(x)\big) = 0$$

holds for all x in the interval I? If there is such a function $y(x)$, we can try to find its derivative at $x = a$ by differentiating the equation $F(x, y) = 0$ implicitly with respect to x, and evaluating the result at (a, b):

$$F_1(x,y) + F_2(x,y)\frac{dy}{dx} = 0,$$

so that

$$\left.\frac{dy}{dx}\right|_{x=a} = -\frac{F_1(a,b)}{F_2(a,b)}, \qquad \text{provided} \qquad F_2(a,b) \neq 0.$$

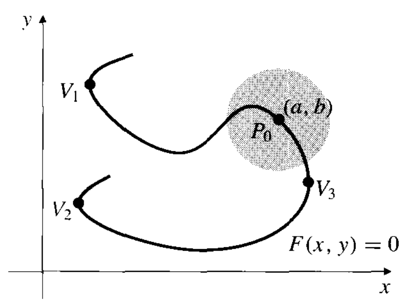

Figure 12.28 The equation $F(x, y) = 0$ can be solved for y as a function of x near P_0 or near any other point except the three points where the curve has a vertical tangent

Observe, however, that the condition $F_2(a, b) \neq 0$ required for the calculation of $y'(a)$ will itself guarantee that the solution $y(x)$ exists. This condition, together with the differentiability of $F(x, y)$ near (a, b), implies that the level curve $F(x, y) = F(a, b)$ has *nonvertical* tangent lines near (a, b), so some part of the level curve near (a, b) must be the graph of a function of x. (See Figure 12.28; the part of the curve $F(x, y) = 0$ in the shaded disk centred at $P_0 = (a, b)$ is the graph of a function $y(x)$ because vertical lines meet that part of the curve only once. The only points on the curve where a disk with that property cannot be drawn are the three points V_1, V_2, and V_3 where the curve has a vertical tangent, that is, where $F_2(x, y) = 0$.) This is a special case of the Implicit Function Theorem, which we will state more generally later in this section.

A similar situation holds for equations involving several variables. We can, for example, ask whether the equation

$$F(x, y, z) = 0$$

defines z as a function of x and y (say, $z = z(x, y)$) near some point P_0 with coordinates (x_0, y_0, z_0) satisfying the equation. If so, and if F has continuous first partials near P_0, then the partial derivatives of z can be found at (x_0, y_0) by implicit differentiation of the equation $F(x, y, z) = 0$ with respect to x and y:

$$F_1(x, y, z) + F_3(x, y, z)\frac{\partial z}{\partial x} = 0 \quad \text{and} \quad F_2(x, y, z) + F_3(x, y, z)\frac{\partial z}{\partial y} = 0,$$

so that

$$\left.\frac{\partial z}{\partial x}\right|_{(x_0,y_0)} = -\frac{F_1(x_0, y_0, z_0)}{F_3(x_0, y_0, z_0)} \quad \text{and} \quad \left.\frac{\partial z}{\partial y}\right|_{(x_0,y_0)} = -\frac{F_2(x_0, y_0, z_0)}{F_3(x_0, y_0, z_0)},$$

provided $F_3(x_0, y_0, z_0) \neq 0$. Since F_3 is the z component of the gradient of F, this condition implies that the level surface of F through P_0 does not have a horizontal normal vector, so it is not vertical (i.e., it is not parallel to the z-axis). Therefore, part of the surface near P_0 must indeed be the graph of a function $z = z(x, y)$. Similarly, $F(x, y, z) = 0$ can be solved for x as a function of y and z near points where $F_1 \neq 0$ and for $y = y(x, z)$ near points where $F_2 \neq 0$.

Example 1 Near what points on the sphere $x^2 + y^2 + z^2 = 1$ can the equation of the sphere be solved for z as a function of x and y? Find $\partial z/\partial x$ and $\partial z/\partial y$ at such points.

Solution The sphere is the level surface $F(x, y, z) = 0$ of the function

$$F(x, y, z) = x^2 + y^2 + z^2 - 1.$$

The above equation can be solved for $z = z(x, y)$ near $P_0 = (x_0, y_0, z_0)$, provided that P_0 is not on the *equator* of the sphere, that is, the circle $x^2 + y^2 = 1$, $z = 0$. The equator consists of those points that satisfy $F_3(x, y, z) = 0$. If P_0 is not on the equator, then it is on either the upper or the lower hemisphere. The upper hemisphere has equation $z = z(x, y) = \sqrt{1 - x^2 - y^2}$, and the lower hemisphere has equation $z = z(x, y) = -\sqrt{1 - x^2 - y^2}$.

If $z \neq 0$, we can calculate the partial derivatives of the solution $z = z(x, y)$ by implicitly differentiating the equation of the sphere: $x^2 + y^2 + z^2 = 1$:

$$2x + 2z\frac{\partial z}{\partial x} = 0, \qquad \text{so} \qquad \frac{\partial z}{\partial x} = -\frac{x}{z},$$

$$2y + 2z\frac{\partial z}{\partial y} = 0, \qquad \text{so} \qquad \frac{\partial z}{\partial y} = -\frac{y}{z}.$$

■

Systems of Equations

Experience with linear equations shows us that systems of such equations can generally be solved for as many variables as there are equations in the system. We would expect, therefore, that a pair of equations in several variables might determine two of those variables as functions of the remaining ones. For instance, we might expect the equations

$$\begin{cases} F(x, y, z, w) = 0 \\ G(x, y, z, w) = 0 \end{cases}$$

to possess, near some point that satisfies them, solutions of one or more of the forms

$$\begin{cases} x = x(z, w) \\ y = y(z, w), \end{cases} \qquad \begin{cases} x = x(y, w) \\ z = z(y, w), \end{cases} \qquad \begin{cases} x = x(y, z) \\ w = w(y, z), \end{cases}$$

$$\begin{cases} y = y(x, w) \\ z = z(x, w), \end{cases} \qquad \begin{cases} y = y(x, z) \\ w = w(x, z), \end{cases} \qquad \begin{cases} z = z(x, y) \\ w = w(x, y). \end{cases}$$

Where such solutions exist, we should be able to differentiate the given system of equations implicitly to find partial derivatives of the solutions.

If you are given a single equation $F(x, y, z) = 0$ and asked to find $\partial x/\partial z$, you would understand that x is intended to be a function of the remaining variables y and z, so there would be no chance of misinterpreting which variable is to be held constant in calculating the partial derivative. Suppose, however, that you are asked to calculate $\partial x/\partial z$ given the system $F(x, y, z, w) = 0$, $G(x, y, z, w) = 0$. The question implies that x is one of the dependent variables and z is one of the independent variables, but does not imply which of y and w is the other dependent variable and which is the other independent variable. In short, which of the situations

$$\begin{cases} x = x(z, w) \\ y = y(z, w) \end{cases} \quad \text{and} \quad \begin{cases} x = x(y, z) \\ w = w(y, z) \end{cases}$$

are we dealing with? As it stands, the question is ambiguous. To avoid this ambiguity we can specify *in the notation for the partial derivative* which variable is to be regarded as the other independent variable and therefore *held fixed* during the differentiation. Thus,

$$\left(\frac{\partial x}{\partial z}\right)_w \quad \text{implies the interpretation} \quad \begin{cases} x = x(z, w) \\ y = y(z, w), \end{cases}$$

$$\left(\frac{\partial x}{\partial z}\right)_y \quad \text{implies the interpretation} \quad \begin{cases} x = x(y, z) \\ w = w(y, z). \end{cases}$$

Example 2 Given the equations $F(x, y, z, w) = 0$ and $G(x, y, z, w) = 0$, where F and G have continuous first partial derivatives, calculate $(\partial x/\partial z)_w$.

Solution We differentiate the two equations with respect to z, regarding x and y as functions of z and w, and holding w fixed:

$$F_1 \frac{\partial x}{\partial z} + F_2 \frac{\partial y}{\partial z} + F_3 = 0$$

$$G_1 \frac{\partial x}{\partial z} + G_2 \frac{\partial y}{\partial z} + G_3 = 0$$

(Note that the terms $F_4(\partial w/\partial z)$ and $G_4(\partial w/\partial z)$ are not present because w and z are independent variables, and w is being held fixed during the differentiation.) The pair of equations above is linear in $\partial x/\partial z$ and $\partial y/\partial z$. Eliminating $\partial y/\partial z$ (or using Cramer's Rule, Theorem 6 of Section 10.6), we obtain

$$\left(\frac{\partial x}{\partial z}\right)_w = -\frac{F_3G_2 - F_2G_3}{F_1G_2 - F_2G_1}.$$

In the light of the examples considered above, you should not be too surprised to learn that the nonvanishing of the denominator $F_1G_2 - F_2G_1$ at some point $P_0 = (x_0, y_0, z_0, w_0)$ satisfying the system $F = 0$, $G = 0$ is sufficient to guarantee that the system does indeed have a solution of the form $x = x(z, w)$, $y = y(z, w)$ near P_0. We will not, however, attempt to prove this fact here. ■

Example 3 Let x, y, u, and v be related by the equations

$$\begin{cases} u = x^2 + xy - y^2 \\ v = 2xy + y^2 \end{cases}$$

Find (a) $(\partial x/\partial u)_v$ and (b) $(\partial x/\partial u)_y$ at the point where $x = 2$ and $y = -1$.

Solution

(a) To calculate $(\partial x/\partial u)_v$ we regard x and y as functions of u and v and differentiate the given equations with respect to u, holding v constant:

$$\begin{aligned} 1 &= \frac{\partial u}{\partial u} = (2x + y)\frac{\partial x}{\partial u} + (x - 2y)\frac{\partial y}{\partial u} \\ 0 &= \frac{\partial v}{\partial u} = 2y\frac{\partial x}{\partial u} + (2x + 2y)\frac{\partial y}{\partial u} \end{aligned}$$

At $x = 2$, $y = -1$ we have

$$\begin{aligned} 1 &= 3\frac{\partial x}{\partial u} + 4\frac{\partial y}{\partial u} \\ 0 &= -2\frac{\partial x}{\partial u} + 2\frac{\partial y}{\partial u}. \end{aligned}$$

Eliminating $\partial y/\partial u$ leads to the result $(\partial x/\partial u)_v = 1/7$.

(b) To calculate $(\partial x/\partial u)_y$ we regard x and v as functions of y and u and differentiate the given equations with respect to u, holding y constant:

$$1 = \frac{\partial u}{\partial u} = (2x + y)\frac{\partial x}{\partial u}, \qquad \frac{\partial v}{\partial u} = 2y\frac{\partial x}{\partial u}.$$

At $x = 2$, $y = -1$ the first equation immediately gives $(\partial x/\partial u)_y = 1/3$. ■

It often happens that the independent variables in a problem are either clear from the context or can be chosen at the outset. In either case, ambiguity is not likely to occur, and we can safely omit subscripts showing which variables are being held constant. The following example is taken from the thermodynamics of an ideal gas.

Example 4 **(Reversible changes in an ideal gas)** The state of a closed system containing n moles of an ideal gas is characterized by three state variables—pressure, volume, and temperature (P, V, and T, respectively)—which satisfy the equation of state for an ideal gas:

$$PV = nRT,$$

where R is a universal constant. This equation can be regarded as defining one of the three variables as a function of the other two. The internal energy, E, and the entropy, S, of the system are thermodynamic quantities that depend on the state variables and hence may be expressed as functions of any two of P, V, and T. Let us choose V and T for the independent variables and so write $E = E(V, T)$ and $S = S(V, T)$. For reversible processes in the system, the first and second laws of thermodynamics imply that infinitesimal changes in these quantities satisfy the differential equation

$$T\,dS = dE + P\,dV.$$

Deduce that, for such processes, E is independent of V and depends only on the temperature T.

Solution We calculate the differentials dS and dE and substitute them into the differential equation to obtain

$$T\left(\frac{\partial S}{\partial V}\,dV + \frac{\partial S}{\partial T}\,dT\right) = \frac{\partial E}{\partial V}\,dV + \frac{\partial E}{\partial T}\,dT + P\,dV.$$

Divide by T, substitute nR/V for P/T (from the equation of state), and collect coefficients of dV and dT on opposite sides of the equation to get

$$\left(\frac{\partial S}{\partial V} - \frac{1}{T}\frac{\partial E}{\partial V} - \frac{nR}{V}\right)dV = \left(\frac{1}{T}\frac{\partial E}{\partial T} - \frac{\partial S}{\partial T}\right)dT.$$

Since dV and dT are independent variables, both coefficients must vanish. Hence

$$\frac{\partial S}{\partial V} = \frac{1}{T}\frac{\partial E}{\partial V} + \frac{nR}{V}$$
$$\frac{\partial S}{\partial T} = \frac{1}{T}\frac{\partial E}{\partial T}.$$

Now differentiate the first of these equations with respect to T and the second with respect to V. Using equality of mixed partials for both S and E, we obtain the desired result:

$$\frac{\partial}{\partial T}\left(\frac{1}{T}\frac{\partial E}{\partial V} + \frac{nR}{V}\right) = \frac{\partial^2 S}{\partial T\partial V} = \frac{\partial^2 S}{\partial V\partial T} = \frac{\partial}{\partial V}\left(\frac{1}{T}\frac{\partial E}{\partial T}\right)$$
$$\frac{-1}{T^2}\frac{\partial E}{\partial V} + \frac{1}{T}\frac{\partial^2 E}{\partial T\partial V} = \frac{1}{T}\frac{\partial^2 E}{\partial V\partial T}$$
$$\frac{-1}{T^2}\frac{\partial E}{\partial V} = 0.$$

It follows that $\partial E/\partial V = 0$, so E is independent of V. ■

Jacobian Determinants

Partial derivatives obtained by implicit differentiation of systems of equations are fractions, the numerators and denominators of which are conveniently expressed in terms of certain determinants called *Jacobians*.

DEFINITION 8

The **Jacobian determinant** (or simply **the Jacobian**) of the two functions, $u = u(x, y)$ and $v = v(x, y)$, with respect to two variables, x and y, is the determinant

$$\frac{\partial(u, v)}{\partial(x, y)} = \begin{vmatrix} \dfrac{\partial u}{\partial x} & \dfrac{\partial u}{\partial y} \\ \dfrac{\partial v}{\partial x} & \dfrac{\partial v}{\partial y} \end{vmatrix}.$$

Similarly, the Jacobian of two functions, $F(x, y, \ldots)$ and $G(x, y, \ldots)$, with respect to the variables, x and y, is the determinant

$$\frac{\partial(F, G)}{\partial(x, y)} = \begin{vmatrix} \dfrac{\partial F}{\partial x} & \dfrac{\partial F}{\partial y} \\ \dfrac{\partial G}{\partial x} & \dfrac{\partial G}{\partial y} \end{vmatrix} = \begin{vmatrix} F_1 & F_2 \\ G_1 & G_2 \end{vmatrix}.$$

The definition above can be extended in the obvious way to give the Jacobian of n functions (or variables) with respect to n variables. For example, the Jacobian of three functions, F, G, and H, with respect to three variables, x, y, and z, is the determinant

$$\frac{\partial(F, G, H)}{\partial(x, y, z)} = \begin{vmatrix} F_1 & F_2 & F_3 \\ G_1 & G_2 & G_3 \\ H_1 & H_2 & H_3 \end{vmatrix}.$$

Jacobians are the determinants of the square Jacobian matrices corresponding to transformations of $\mathbb{R}^n$ to $\mathbb{R}^n$ as discussed briefly in Section 12.6.

Example 5 In terms of Jacobians, the value of $(\partial x/\partial z)_w$, obtained from the system of equations

$$F(x, y, z, w) = 0, \qquad G(x, y, z, w) = 0$$

in Example 2 can be expressed in the form

$$\left(\frac{\partial x}{\partial z}\right)_w = -\frac{\dfrac{\partial(F, G)}{\partial(z, y)}}{\dfrac{\partial(F, G)}{\partial(x, y)}}.$$

Observe the pattern here. The denominator is the Jacobian of F and G with respect to the two *dependent* variables, x and y. The numerator is the same Jacobian except that the dependent variable x is replaced by the independent variable z. ■

The pattern observed above is general. We state it formally in the Implicit Function Theorem below.

The Implicit Function Theorem

The Implicit Function Theorem guarantees that systems of equations can be solved for certain variables as functions of other variables under certain circumstances, and provides formulas for the partial derivatives of the solution functions. Before stating it, we consider a simple illustrative example.

Example 6 Consider the system of linear equations

$$\begin{aligned} F(x, y, s, t) &= a_1x + b_1y + c_1s + d_1t + e_1 = 0 \\ G(x, y, s, t) &= a_2x + b_2y + c_2s + d_2t + e_2 = 0. \end{aligned}$$

This system can be written in matrix form:

$$\mathcal{A}\begin{pmatrix} x \\ y \end{pmatrix} + \mathcal{C}\begin{pmatrix} s \\ t \end{pmatrix} + \mathcal{E} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

where

$$\mathcal{A} = \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}, \qquad \mathcal{C} = \begin{pmatrix} c_1 & d_1 \\ c_2 & d_2 \end{pmatrix}, \quad \text{and} \quad \mathcal{E} = \begin{pmatrix} e_1 \\ e_2 \end{pmatrix}.$$

The equations can be solved for x and y as functions of s and t provided $\det(\mathcal{A}) \neq 0$, for this implies the existence of the inverse matrix $\mathcal{A}^{-1}$ (Theorem 4 of Section 10.6), so

$$\begin{pmatrix} x \\ y \end{pmatrix} = -\mathcal{A}^{-1}\left(\mathcal{C}\begin{pmatrix} s \\ t \end{pmatrix} + \mathcal{E}\right).$$

Observe that $\det(\mathcal{A}) = \partial(F, G)/\partial(x, y)$, so the nonvanishing of this Jacobian guarantees that the equations can be solved for x and y. ■

THEOREM 8

The Implicit Function Theorem

Consider a system of n equations in $n + m$ variables,

$$\begin{cases} F_{(1)}(x_1, x_2, \dots, x_m, y_1, y_2, \dots, y_n) = 0 \\ F_{(2)}(x_1, x_2, \dots, x_m, y_1, y_2, \dots, y_n) = 0 \\ \quad\vdots \\ F_{(n)}(x_1, x_2, \dots, x_m, y_1, y_2, \dots, y_n) = 0, \end{cases}$$

and a point $P_0 = (a_1, a_2, \dots, a_m, b_1, b_2, \dots, b_n)$ that satisfies the system. Suppose each of the functions $F_{(i)}$ has continuous first partial derivatives with respect to each of the variables x_j and y_k, $(i = 1, \dots, n,\ j = 1, \dots, m,\ k = 1, \dots, n)$, near P_0. Finally, suppose that

$$\left.\frac{\partial(F_{(1)}, F_{(2)}, \dots, F_{(n)})}{\partial(y_1, y_2, \dots, y_n)}\right|_{P_0} \neq 0.$$

Then the system can be solved for $y_1, y_2, \ldots, y_n$ as functions of $x_1, x_2, \ldots, x_m$ near P_0. That is, there exist functions

$$\phi_j(a_1, \ldots, a_m) = b_j, \quad (j = 1, \ldots, n),$$

and such that the equations

$$F_{(1)}\Big(x_1, \ldots, x_m, \phi_1(x_1, \ldots, x_m), \ldots, \phi_n(x_1, \ldots, x_m)\Big) = 0,$$
$$F_{(2)}\Big(x_1, \ldots, x_m, \phi_1(x_1, \ldots, x_m), \ldots, \phi_n(x_1, \ldots, x_m)\Big) = 0,$$
$$\vdots$$
$$F_{(n)}\Big(x_1, \ldots, x_m, \phi_1(x_1, \ldots, x_m), \ldots, \phi_n(x_1, \ldots, x_m)\Big) = 0,$$

hold for all $(x_1, \ldots, x_m)$ sufficiently near $(a_1, \ldots, a_m)$.

Moreover,

$$\frac{\partial \phi_i}{\partial x_j} = \left(\frac{\partial y_i}{\partial x_j}\right)_{x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_m} = -\frac{\dfrac{\partial(F_{(1)}, F_{(2)}, \ldots, F_{(n)})}{\partial(y_1, \ldots, x_j, \ldots, y_n)}}{\dfrac{\partial(F_{(1)}, F_{(2)}, \ldots, F_{(n)})}{\partial(y_1, \ldots, y_i, \ldots, y_n)}}.$$

Remark The formula for the partial derivatives is a consequence of Cramer's Rule (Theorem 6 of Section 10.6) applied to the n linear equations in the n unknowns $\partial y_1/\partial x_j, \ldots, \partial y_n/\partial x_j$ obtained by differentiating each of the equations in the given system with respect to x_j.

Example 7 Show that the system

$$\begin{cases} xy^2 + xzu + yv^2 = 3 \\ x^3yz + 2xv - u^2v^2 = 2 \end{cases}$$

can be solved for (u, v) as a (vector) function of (x, y, z) near the point P_0 where $(x, y, z, u, v) = (1, 1, 1, 1, 1)$, and find the value of $\partial v/\partial y$ for the solution at $(x, y, z) = (1, 1, 1)$.

Solution Let $\begin{cases} F(x, y, z, u, v) = xy^2 + xzu + yv^2 - 3 \\ G(x, y, z, u, v) = x^3yz + 2xv - u^2v^2 - 2 \end{cases}$. Then

$$\left.\frac{\partial(F, G)}{\partial(u, v)}\right|_{P_0} = \left.\begin{vmatrix} xz & 2yv \\ -2uv^2 & 2x - 2u^2v \end{vmatrix}\right|_{P_0} = \begin{vmatrix} 1 & 2 \\ -2 & 0 \end{vmatrix} = 4.$$

Since this Jacobian is not zero, the Implicit Function Theorem assures us that the given equations can be solved for u and v as functions of x, y, and z, that is, for $(u, v) = \mathbf{f}(x, y, z)$. Since

$$\left.\frac{\partial(F,G)}{\partial(u,y)}\right|_{P_0} = \left|\begin{matrix} xz & 2xy+v^2 \\ -2uv^2 & x^3z \end{matrix}\right|_{P_0} = \left|\begin{matrix} 1 & 3 \\ -2 & 1 \end{matrix}\right| = 7,$$

we have

$$\left(\frac{\partial v}{\partial y}\right)_{x,z} = -\left.\frac{\dfrac{\partial(F,G)}{\partial(u,y)}}{\dfrac{\partial(F,G)}{\partial(u,v)}}\right|_{P_0} = -\frac{7}{4}.$$

Remark If all we wanted in this example was to calculate $\partial v/\partial y$, it would have been easier to use the technique of Example 3 and differentiate the given equations directly with respect to y, holding x and z fixed.

Example 8 If the equations $x = u^2 + v^2$ and $y = uv$ are solved for u and v in terms of x and y, find, where possible,

$$\frac{\partial u}{\partial x}, \quad \frac{\partial u}{\partial y}, \quad \frac{\partial v}{\partial x}, \quad \text{and} \quad \frac{\partial v}{\partial y}.$$

Hence, show that $\dfrac{\partial(u,v)}{\partial(x,y)} = 1\Big/\dfrac{\partial(x,y)}{\partial(u,v)}$, provided the denominator is not zero.

Solution The given equations can be rewritten in the form

$$F(u,v,x,y) = u^2 + v^2 - x = 0$$
$$G(u,v,x,y) = uv - y = 0.$$

Let

$$J = \frac{\partial(F,G)}{\partial(u,v)} = \begin{vmatrix} 2u & 2v \\ v & u \end{vmatrix} = 2(u^2 - v^2) = \frac{\partial(x,y)}{\partial(u,v)}.$$

If $u^2 \neq v^2$, then $J \neq 0$ and we can calculate the required partial derivatives:

$$\frac{\partial u}{\partial x} = -\frac{1}{J}\frac{\partial(F,G)}{\partial(x,v)} = -\frac{1}{J}\begin{vmatrix} -1 & 2v \\ 0 & u \end{vmatrix} = \frac{u}{2(u^2-v^2)}$$
$$\frac{\partial u}{\partial y} = -\frac{1}{J}\frac{\partial(F,G)}{\partial(y,v)} = -\frac{1}{J}\begin{vmatrix} 0 & 2v \\ -1 & u \end{vmatrix} = \frac{-2v}{2(u^2-v^2)}$$
$$\frac{\partial v}{\partial x} = -\frac{1}{J}\frac{\partial(F,G)}{\partial(u,x)} = -\frac{1}{J}\begin{vmatrix} 2u & -1 \\ v & 0 \end{vmatrix} = \frac{-v}{2(u^2-v^2)}$$
$$\frac{\partial v}{\partial y} = -\frac{1}{J}\frac{\partial(F,G)}{\partial(u,y)} = -\frac{1}{J}\begin{vmatrix} 2u & 0 \\ v & -1 \end{vmatrix} = \frac{2u}{2(u^2-v^2)}.$$

Thus,

$$\frac{\partial(u,v)}{\partial(x,y)} = \frac{1}{J^2}\begin{vmatrix} u & -2v \\ -v & 2u \end{vmatrix} = \frac{J}{J^2} = \frac{1}{J} = \frac{1}{\dfrac{\partial(x,y)}{\partial(u,v)}}.$$

Remark Note in the above example that $\partial u/\partial x \neq 1/(\partial x/\partial u)$. This should be contrasted with the single-variable situation where, if $y = f(x)$ and $dy/dx \neq 0$, then $x = f^{-1}(y)$ and $dx/dy = 1/(dy/dx)$. This is another reason for distinguishing between "∂" and "d." It is the Jacobian rather than any single partial derivative that takes the place of the ordinary derivative in such situations.

Remark Let us look briefly at the general case of invertible transformations from $\mathbb{R}^n$ to $\mathbb{R}^n$. Suppose that $\mathbf{y} = \mathbf{f}(\mathbf{x})$ and $\mathbf{z} = \mathbf{g}(\mathbf{y})$ are both functions from $\mathbb{R}^n$ to $\mathbb{R}^n$ whose components have continuous first partial derivatives. As shown in Section 12.6, the Chain Rule implies that

$$\begin{pmatrix} \frac{\partial z_1}{\partial x_1} & \cdots & \frac{\partial z_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial z_n}{\partial x_1} & \cdots & \frac{\partial z_n}{\partial x_n} \end{pmatrix} = \begin{pmatrix} \frac{\partial z_1}{\partial y_1} & \cdots & \frac{\partial z_1}{\partial y_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial z_n}{\partial y_1} & \cdots & \frac{\partial z_n}{\partial y_n} \end{pmatrix} \begin{pmatrix} \frac{\partial y_1}{\partial x_1} & \cdots & \frac{\partial y_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial y_n}{\partial x_1} & \cdots & \frac{\partial y_n}{\partial x_n} \end{pmatrix}.$$

This is just the Chain Rule for the composition $\mathbf{z} = \mathbf{g}\big(\mathbf{f}(\mathbf{x})\big)$. It follows from Theorem 3(b) of Section 10.6 that the determinants of these matrices satisfy a similar equation:

$$\frac{\partial(z_1 \cdots z_n)}{\partial(x_1 \cdots x_n)} = \frac{\partial(z_1 \cdots z_n)}{\partial(y_1 \cdots y_n)} \frac{\partial(y_1 \cdots y_n)}{\partial(x_1 \cdots x_n)}.$$

If $\mathbf{f}$ is one-to-one and $\mathbf{g}$ is the inverse of $\mathbf{f}$, then $\mathbf{z} = \mathbf{g}\big(\mathbf{f}(\mathbf{x})\big) = \mathbf{x}$, and $\partial(z_1 \cdots z_n)/\partial(x_1 \cdots x_n) = 1$, the determinant of the identity matrix. Thus

$$\frac{\partial(x_1 \cdots x_n)}{\partial(y_1 \cdots y_n)} = \frac{1}{\dfrac{\partial(y_1 \cdots y_n)}{\partial(x_1 \cdots x_n)}}.$$

In fact, the nonvanishing of either of these determinants is sufficient to guarantee that $\mathbf{f}$ is one-to-one and has an inverse. This is a special case of the Implicit Function Theorem.

We will encounter Jacobians again when we study transformations of coordinates in multiple integrals in Chapter 14.

Exercises 12.8

In Exercises 1–12, calculate the indicated derivative from the given equation(s). What condition on the variables will guarantee the existence of a solution that has the indicated derivative? Assume that any general functions F, G, and H have continuous first partial derivatives.

1. $\dfrac{dx}{dy}$ if $xy^3 + x^4 y = 2$

2. $\dfrac{\partial x}{\partial y}$ if $xy^3 = y - z$

3. $\dfrac{\partial z}{\partial y}$ if $z^2 + xy^3 = \dfrac{xz}{y}$

4. $\dfrac{\partial y}{\partial z}$ if $e^{yz} - x^2 z \ln y = \pi$

5. $\dfrac{\partial x}{\partial w}$ if $x^2y^2 + y^2z^2 + z^2t^2 + t^2w^2 - xw = 0$

6. $\dfrac{dy}{dx}$ if $F(x, y, x^2 - y^2) = 0$

7. $\dfrac{\partial u}{\partial x}$ if $G(x, y, z, u, v) = 0$

8. $\dfrac{\partial z}{\partial x}$ if $F(x^2 - z^2, y^2 + xz) = 0$

9. $\dfrac{\partial w}{\partial t}$ if $H(u^2 w, v^2 t, wt) = 0$

10. $\left(\dfrac{\partial y}{\partial x}\right)_u$ if $xyuv = 1$ and $x + y + u + v = 0$

11. $\left(\frac{\partial x}{\partial y}\right)_z$ if $x^2 + y^2 + z^2 + w^2 = 1$, and $x + 2y + 3z + 4w = 2$

12. $\frac{du}{dx}$ if $x^2 y + y^2 u - u^3 = 0$ and $x^2 + yu = 1$

13. If $x = u^3 + v^3$ and $y = uv - v^2$ are solved for u and v in terms of x and y, evaluate

$$\frac{\partial u}{\partial x}, \quad \frac{\partial u}{\partial y}, \quad \frac{\partial v}{\partial x}, \quad \frac{\partial v}{\partial y}, \quad \text{and} \quad \frac{\partial(u, v)}{\partial(x, y)}$$

at the point where $u = 1$ and $v = 1$.

14. Near what points (r, s) can the transformation

$$x = r^2 + 2s, \qquad y = s^2 - 2r$$

be solved for r and s as functions of x and y? Calculate the values of the first partial derivatives of the solution at the origin.

15. Evaluate the Jacobian $\partial(x, y)/\partial(r, \theta)$ for the transformation to polar coordinates: $x = r\cos\theta$, $y = r\sin\theta$. Near what points (r, θ) is the transformation one-to-one and therefore invertible to give r and θ as functions of x and y?

16. Evaluate the Jacobian $\partial(x, y, z)/\partial(\rho, \phi, \theta)$, where

$$x = \rho \sin\phi \cos\theta, \quad y = \rho \sin\phi \sin\theta, \quad \text{and } z = \rho\cos\phi.$$

This is the transformation from Cartesian to spherical polar coordinates in 3-space that we will consider in Section 14.6. Near what points is the transformation one-to-one and hence invertible to give ρ, ϕ, and θ as functions of x, y, and z?

17. Show that the equations

$$\begin{cases} xy^2 + zu + v^2 = 3 \\ x^3 z + 2y - uv = 2 \\ xu + yv - xyz = 1 \end{cases}$$

can be solved for x, y, and z as functions of u and v near the point P_0 where $(x, y, z, u, v) = (1, 1, 1, 1, 1)$, and find $(\partial y/\partial u)_v$ at $(u, v) = (1, 1)$.

18. Show that the equations $\begin{cases} xe^y + uz - \cos v = 2 \\ u\cos y + x^2 v - yz^2 = 1 \end{cases}$ can be solved for u and v as functions of x, y, and z near the point P_0 where $(x, y, z) = (2, 0, 1)$ and $(u, v) = (1, 0)$, and find $(\partial u/\partial z)_{x,y}$ at $(x, y, z) = (2, 0, 1)$.

19. Find dx/dy from the system

$$F(x, y, z, w) = 0, \quad G(x, y, z, w) = 0, \quad H(x, y, z, w) = 0.$$

20. Given the system

$$F(x, y, z, u, v) = 0$$
$$G(x, y, z, u, v) = 0$$
$$H(x, y, z, u, v) = 0,$$

how many possible interpretations are there for $\partial x/\partial y$? Evaluate them.

21. Given the system

$$F(x_1, x_2, \ldots, x_8) = 0$$
$$G(x_1, x_2, \ldots, x_8) = 0$$
$$H(x_1, x_2, \ldots, x_8) = 0,$$

how many possible interpretations are there for the partial $\partial x_1/\partial x_2$? Evaluate $\left(\partial x_1/\partial x_2\right)_{x_4, x_6, x_7, x_8}$.

22. If $F(x, y, z) = 0$ determines z as a function of x and y, calculate $\partial^2 z/\partial x^2$, $\partial^2 z/\partial x\partial y$, and $\partial^2 z/\partial y^2$ in terms of the partial derivatives of F.

23. If $x = u + v$, $y = uv$, and $z = u^2 + v^2$ define z as a function of x and y, find $\partial z/\partial x$, $\partial z/\partial y$, and $\partial^2 z/\partial x\partial y$.

24. A certain gas satisfies the law $pV = T - \frac{4p}{T^2}$, where p = pressure, V = volume, and T = temperature.
 (a) Calculate $\partial T/\partial p$ and $\partial T/\partial V$ at the point where $p = V = 1$ and $T = 2$.
 (b) If measurements of p and V yield the values $p = 1 \pm 0.001$ and $V = 1 \pm 0.002$, find the approximate maximum error in the calculated value $T = 2$.

25. If $F(x, y, z) = 0$, show that $\left(\frac{\partial x}{\partial y}\right)_z \left(\frac{\partial y}{\partial z}\right)_x \left(\frac{\partial z}{\partial x}\right)_y = -1.$ Derive analogous results for $F(x, y, z, u) = 0$ and for $F(x, y, z, u, v) = 0$. What is the general case?

* 26. If the equations $F(x, y, u, v) = 0$ and $G(x, y, u, v) = 0$ are solved for x and y as functions of u and v, show that

$$\frac{\partial(x, y)}{\partial(u, v)} = \frac{\partial(F, G)}{\partial(u, v)} \Bigg/ \frac{\partial(F, G)}{\partial(x, y)}.$$

* 27. If the equations $x = f(u, v)$, $y = g(u, v)$ can be solved for u and v in terms of x and y, show that

$$\frac{\partial(u, v)}{\partial(x, y)} = 1 \Bigg/ \frac{\partial(x, y)}{\partial(u, v)}.$$

Hint: use the result of Exercise 26.

* 28. If $x = f(u, v)$, $y = g(u, v)$, $u = h(r, s)$, and $v = k(r, s)$, then x and y can be expressed as functions of r and s. Verify by direct calculation that

$$\frac{\partial(x, y)}{\partial(r, s)} = \frac{\partial(x, y)}{\partial(u, v)} \frac{\partial(u, v)}{\partial(r, s)}.$$

This is a special case of the Chain Rule for Jacobians.

* **29.** Two functions, $f(x, y)$ and $g(x, y)$, are said to be functionally dependent if one is a function of the other; that is, if there exists a single-variable function $k(t)$ such that

* **30.** Prove the converse of the previous exercise as follows: Let $u = f(x, y)$ and $v = g(x, y)$, and suppose that $\partial(u, v)/\partial(x, y) = \partial(f, g)/\partial(x, y)$ is identically zero for all

12.9 Taylor Series and Approximations

As is the case for functions of one variable, power series representations and their partial sums (Taylor polynomials) can provide an efficient method for determining the behaviour of a smooth function of several variables near a point in its domain. In this section we will look briefly at the extension of Taylor series to such functions. As usual, we will develop the machinery for functions of two variables, but the extension to more variables should be clear.

As a starting point, recall Taylor's Formula for a function $F(x)$ with continuous derivatives of order up to $n + 1$ on the interval $[a, a + h]$:

$$F(a+h) = F(a) + F'(a)h + \frac{F''(a)}{2!}h^2 + \cdots + \frac{F^{(n)}(a)}{n!}h^n + \frac{F^{(n+1)}(X)}{(n+1)!}h^{n+1},$$

where X is some number between a and $a + h$. (The last term in the formula is the *Lagrange* form of the remainder.) In the special case where $a = 0$ and $h = 1$, this formula becomes

$$F(1) = F(0) + F'(0) + \frac{F''(0)}{2!} + \cdots + \frac{F^{(n)}(0)}{n!} + \frac{F^{(n+1)}(\theta)}{(n+1)!}$$

for some θ between 0 and 1.

Now suppose that $f(x, y)$ has continuous partial derivatives up to order $n + 1$ at all points in an open set containing the line segment joining the points (a, b) and $(a + h, b + k)$ in its domain. We can find the Taylor Formula for $f(a + h, b + k)$ in powers of h and k by applying the one-variable formula above to the function

$$F(t) = f(a + th, b + tk), \qquad 0 \le t \le 1.$$

Clearly, $F(0) = f(a, b)$ and $F(1) = f(a + h, b + k)$. Let us calculate some derivatives of F:

$$\begin{aligned}
F'(t) &= hf_1(a + th, b + tk) + kf_2(a + th, b + tk),\\
F''(t) &= h^2 f_{11}(a + th, b + tk) + 2hkf_{12}(a + th, b + tk)\\
&\quad + k^2 f_{22}(a + th, b + tk),\\
F'''(t) &= \left(h^3 f_{111} + 3h^2 k f_{112} + 3hk^2 f_{122} + k^3 f_{222}\right)\Big|_{(a+th,b+tk)}.
\end{aligned}$$

The pattern of binomial coefficients is pretty obvious here, but the notation, involving subscripts to denote partial derivatives of f, becomes more and more unwieldy as the order of the derivatives increases. The notation can be simplified greatly by using $D_1 f$ and $D_2 f$ to denote the first partials of f with respect to its first and second variables. Since h and k are constant and mixed partials commute ($D_1 D_2 f = D_2 D_1 f$), we have

$$h^2 D_1^2 f + 2hk D_1 D_2 f + k^2 D_2^2 f = (hD_1 + kD_2)^2 f,$$

and so on. Therefore,

$$\begin{aligned}
F'(t) &= \left(hD_1 + kD_2\right) f(a+th, b+tk),\\
F''(t) &= \left(hD_1 + kD_2\right)^2 f(a+th, b+tk),\\
F'''(t) &= \left(hD_1 + kD_2\right)^3 f(a+th, b+tk),\\
&\ \vdots\\
F^{(m)}(t) &= \left(hD_1 + kD_2\right)^m f(a+th, b+tk).
\end{aligned}$$

In particular, $F^{(m)}(0) = \left(hD_1 + kD_2\right)^m f(a,b)$. Hence, the Taylor Formula for $f(a+h, b+k)$ is

$$\begin{aligned}
f(a+h, b+k) &= \sum_{m=0}^{n} \frac{1}{m!}\left(hD_1 + kD_2\right)^m f(a,b) + R_n(h,k)\\
&= \sum_{m=0}^{n}\sum_{j=0}^{m} C_{mj}\, D_1^j D_2^{m-j} f(a,b)\, h^j k^{m-j} + R_n(h,k),
\end{aligned}$$

where, using the binomial expansion, we have

$$C_{mj} = \frac{1}{m!}\binom{m}{j} = \frac{1}{m!}\,\frac{m!}{j!(m-j)!} = \frac{1}{j!(m-j)!},$$

and where the remainder term is given by

$$R_n(h,k) = \frac{1}{(n+1)!}\left(hD_1 + kD_2\right)^{n+1} f(a+\theta h, b+\theta k)$$

for some θ between 0 and 1. If f has partial derivatives of all orders and

$$\lim_{n\to\infty} R_n(h,k) = 0,$$

then $f(a+h, b+k)$ can be expressed as a Taylor series in powers of h and k:

$$f(a+h, b+k) = \sum_{m=0}^{\infty}\sum_{j=0}^{m} \frac{1}{j!(m-j)!} D_1^j D_2^{m-j} f(a,b)\, h^j k^{m-j}.$$

As for functions of one variable, the Taylor polynomial of degree n,

$$P_n(x,y) = \sum_{m=0}^{n}\sum_{j=0}^{m} \frac{1}{j!(m-j)!} D_1^j D_2^{m-j} f(a,b)\,(x-a)^j (y-b)^{m-j},$$

provides the "best" nth-degree polynomial approximation to $f(x,y)$ near (a,b). For $n = 1$ this approximation reduces to the tangent plane approximation

$$f(x,y) \approx f(a,b) + f_1(a,b)(x-a) + f_2(a,b)(y-b).$$

Example 1 Find a second-degree polynomial approximation to the function $f(x, y) = \sqrt{x^2 + y^3}$ near the point $(1, 2)$ and use it to estimate $\sqrt{(1.02)^2 + (1.97)^3}$.

Solution For the second-degree approximation we need the values of the partial derivatives of f up to second order at (1,2). We have

$$\begin{aligned}
f(x, y) &= \sqrt{x^2 + y^3} & f(1, 2) &= 3 \\
f_1(x, y) &= \frac{x}{\sqrt{x^2 + y^3}} & f_1(1, 2) &= \frac{1}{3} \\
f_2(x, y) &= \frac{3y^2}{2\sqrt{x^2 + y^3}} & f_2(1, 2) &= 2 \\
f_{11}(x, y) &= \frac{y^3}{(x^2 + y^3)^{3/2}} & f_{11}(1, 2) &= \frac{8}{27} \\
f_{12}(x, y) &= \frac{-3xy^2}{2(x^2 + y^3)^{3/2}} & f_{12}(1, 2) &= -\frac{2}{9} \\
f_{22}(x, y) &= \frac{12x^2y + 3y^4}{4(x^2 + y^3)^{3/2}} & f_{22}(1, 2) &= \frac{2}{3}.
\end{aligned}$$

Thus,

$$f(1 + h, 2 + k) \approx 3 + \frac{1}{3}h + 2k + \frac{1}{2!}\left(\frac{8}{27}h^2 + 2\left(-\frac{2}{9}\right)hk + \frac{2}{3}k^2\right)$$

or, setting $x = 1 + h$ and $y = 2 + k$,

$$f(x, y) = 3 + \frac{1}{3}(x - 1) + 2(y - 2) + \frac{4}{27}(x - 1)^2 - \frac{2}{9}(x - 1)(y - 2) + \frac{1}{3}(y - 2)^2.$$

This is the required second-degree Taylor polynomial for f near $(1, 2)$. Therefore,

$$\begin{aligned}
\sqrt{(1.02)^2 + (1.97)^3} &= f(1 + 0.02, 2 - 0.03) \\
&\approx 3 + \frac{1}{3}(0.02) + 2(-0.03) + \frac{4}{27}(0.02)^2 \\
&\quad - \frac{2}{9}(0.02)(-0.03) + \frac{1}{3}(-0.03)^2 \\
&\approx 2.9471593\,.
\end{aligned}$$

(For comparison purposes: the true value is 2.9471636... The approximation is accurate to 6 significant figures.)

■

As observed for functions of one variable, it is not usually necessary to calculate derivatives in order to determine the coefficients in a Taylor series or Taylor polynomial. It is often much easier to perform algebraic manipulations on known series.

For instance, the above example could have been done by writing f in the form

$$\begin{aligned} f(1+h, 2+k) &= \sqrt{(1+h)^2+(2+k)^3} \\ &= \sqrt{9+2h+h^2+12k+6k^2+k^3} \\ &= 3\sqrt{1+\frac{2h+h^2+12k+6k^2+k^3}{9}} \end{aligned}$$

and then applying the binomial expansion

$$\sqrt{1+t} = 1 + \frac{1}{2}t + \frac{1}{2!}\left(\frac{1}{2}\right)\left(-\frac{1}{2}\right)t^2 + \cdots$$

with $t = \dfrac{2h+h^2+12k+6k^2+k^3}{9}$ to obtain the terms up to second degree in h and k.

Example 2 Find the Taylor polynomial of degree 3 in powers of x and y for the function $f(x, y) = e^{x-2y}$.

Solution The required Taylor polynomial will be the Taylor polynomial of degree 3 for e^t evaluated at $t = x - 2y$:

$$\begin{aligned} P_3(x, y) &= 1 + (x-2y) + \frac{1}{2!}(x-2y)^2 + \frac{1}{3!}(x-2y)^3 \\ &= 1 + x - 2y + \frac{1}{2}x^2 - 2xy + 2y^2 + \frac{1}{6}x^3 - x^2y + 2xy^2 - \frac{4}{3}y^3. \end{aligned}$$

■

Remark Maple can, of course, be used to compute multivariate Taylor polynomials via its function **mtaylor**, which must be read in from the Maple library before it can be used, because it is not part of the Maple kernel.

```
> readlib(mtaylor):
```

Arguments fed to `mtaylor` are as follows:

(a) an expression involving the expansion variables

(b) a list whose elements are either variable names or equations of the form `variable=value` giving the coordinates of the point about which the expansion is calculated. (Just naming a variable is equivalent to using the equation `variable=0`.)

(c) (optionally) a positive integer m forcing the order of the computed Taylor polynomial to be less than m. If m is not specified, the value of Maple's global variable "Order" is used. The default value is 6.

A few examples should suffice.

```
> mtaylor(cos(x+y^2),[x,y]);
```

$$1 - \frac{1}{2}x^2 - y^2x + \frac{1}{24}x^4 - \frac{1}{2}y^4 + \frac{1}{6}y^2x^3$$

```
> mtaylor(cos(x+y^2),[x=Pi,y],5);
```

$$-1 + \frac{1}{2}(x-\pi)^2 + y^2(x-\pi) - \frac{1}{24}(x-\pi)^4 + \frac{1}{2}y^4$$

```
> mtaylor(g(x,y),[x=a,y=b],3);
```

$$g(a,b) + D_1(g)(a,b)(x-a) + D_2(g)(a,b)(y-b) + \frac{1}{2}D_{1,1}(g)(a,b)(x-a)^2$$
$$+ (x-a)D_{1,2}(g)(a,b)(y-b) + \frac{1}{2}D_{2,2}(g)(a,b)(y-b)^2$$

In Maple V the function `mtaylor` is a bit quirky. It has a tendency to expand linear terms; for example, in an expansion about $x = 1$ and $y = -2$, it may rewrite terms $2 + (x-1) + 2(y+2)$ in the form $5 + x + 2y$. Also, it will often include terms of higher order than requested when dealing with square roots or other fractional powers. The following example illustrates both of these oddities:

```
> mtaylor(sqrt(x+y),[x=1,y=3],1);
```

$$1 + \frac{1}{4}x + \frac{1}{4}y - \frac{1}{64}(x-1)^2 - \frac{1}{32}(x-1)(y-3) - \frac{1}{64}(y-3)^2$$

Approximating Implicit Functions

In the previous section we saw how to determine whether an equation in several variables could be solved for one of those variables as a function of the others. Even when such a solution is known to exist, it is not usually possible to find an exact formula for it. However, if the equation involves only smooth functions, then the solution will have a Taylor series. We can determine at least the first several coefficients in that series and thus obtain a useful approximation to the solution. The following example shows the technique.

Example 3 Show that the equation

$$\sin(x+y) = xy + 2x$$

has a solution of the form $y = f(x)$ near $x = 0$ satisfying $f(0) = 0$, and find the terms up to fourth degree for the Taylor series for $f(x)$ in powers of x.

Solution The given equation can be written in the form $F(x, y) = 0$, where

$$F(x,y) = \sin(x+y) - xy - 2x.$$

Since $F(0,0) = 0$ and $F_2(0,0) = \cos(0) = 1 \neq 0$, the equation has a solution $y = f(x)$ near $x = 0$ satisfying $f(0) = 0$ by the implicit function theorem. It is not possible to calculate $f(x)$ exactly, but it will have a Maclaurin series of the form

$$y = f(x) = a_1x + a_2x^2 + a_3x^3 + a_4x^4 + \cdots.$$

(There is no constant term because $f(0) = 0$.) We can substitute this series into the given equation and keep track of terms up to degree 4 in order to calculate the coefficients a_1, a_2, a_3, and a_4. For the left side we use the Maclaurin series for sin to obtain

$$\begin{aligned}\sin(x+y) &= \sin\Big((1+a_1)x + a_2x^2 + a_3x^3 + a_4x^4 + \cdots\Big)\\ &= (1+a_1)x + a_2x^2 + a_3x^3 + a_4x^4 + \cdots\\ &\quad - \frac{1}{3!}\Big((1+a_1)x + a_2x^2 + \cdots\Big)^3 + \cdots\\ &= (1+a_1)x + a_2x^2 + \Big(a_3 - \frac{1}{6}(1+a_1)^3\Big)x^3\\ &\quad + \Big(a_4 - \frac{3}{6}(1+a_1)^2a_2\Big)x^4 + \cdots.\end{aligned}$$

The right side is

$$xy + 2x = 2x + a_1x^2 + a_2x^3 + a_3x^4 + \cdots.$$

Equating coefficients of like powers of x, we obtain

$$\begin{aligned} 1 + a_1 &= 2 & a_1 &= 1\\ a_2 &= a_1 & a_2 &= 1\\ a_3 - \frac{1}{6}(1+a_1)^3 &= a_2 & a_3 &= \frac{7}{3}\\ a_4 - \frac{1}{2}(1+a_1)^2a_2 &= a_3 & a_4 &= \frac{13}{3}.\end{aligned}$$

Thus

$$y = f(x) = x + x^2 + \frac{7}{3}x^3 + \frac{13}{3}x^4 + \cdots.$$

(We could have obtained more terms in the series by keeping track of higher powers of x in the substitution process.)

■

Remark From the series for $f(x)$ obtained above, we can determine the values of the first four derivatives of f at $x = 0$. Remember that

$$a_k = \frac{f^{(k)}(0)}{k!}.$$

We have, therefore,

$$\begin{aligned} f'(0) &= a_1 = 1 & f''(0) &= 2!a_2 = 2\\ f'''(0) &= 3!a_3 = 14 & f^{(4)}(0) &= 4!a_4 = 104.\end{aligned}$$

We could have done the example by first calculating these derivatives by implicit differentiation of the given equation and then determining the series coefficients from them. This would have been a much more difficult way to do it. (Try it and see.)

Exercises 12.9

In Exercises 1–6, find the Taylor series for the given function about the indicated point.

1. $f(x, y) = \dfrac{1}{2 + xy^2}$, $(0, 0)$
2. $f(x, y) = \ln(1 + x + y + xy)$, $(0, 0)$
3. $f(x, y) = \tan^{-1}(x + xy)$, $(0, -1)$
4. $f(x, y) = x^2 + xy + y^3$, $(1, -1)$
5. $f(x, y) = e^{x^2+y^2}$, $(0, 0)$
6. $f(x, y) = \sin(2x + 3y)$, $(0, 0)$

In Exercises 7–12, find Taylor polynomials of the indicated degree for the given functions near the given point. After calculating them by hand, try to get the same results using Maple's `mtaylor` function.

7. $f(x, y) = \dfrac{1}{2 + x - 2y}$, degree 3, near $(2, 1)$
8. $f(x, y) = \ln(x^2 + y^2)$, degree 3, near $(1, 0)$
9. $f(x, y) = \displaystyle\int_0^{x+y^2} e^{-t^2}\,dt$, degree 3, near $(0, 0)$
10. $f(x, y) = \cos(x + \sin y)$, degree 4, near $(0, 0)$
11. $f(x, y) = \dfrac{\sin x}{y}$, degree 2, near $(\frac{\pi}{2}, 1)$
12. $f(x, y) = \dfrac{1 + x}{1 + x^2 + y^4}$, degree 2, near $(0, 0)$

In Exercises 13–14, show that, for x near the indicated point $x = a$, the given equation has a solution of the form $y = f(x)$ taking on the indicated value at that point. Find the first three nonzero terms of the Taylor series for $f(x)$ in powers of $x - a$.

* 13. $x \sin y = y + \sin x$, near $x = 0$, with $f(0) = 0$

* 14. $\sqrt{1 + xy} = 1 + x + \ln(1 + y)$, near $x = 0$, with $f(0) = 0$

* 15. Show that the equation $x + 2y + z + e^{2z} = 1$ has a solution of the form $z = f(x, y)$ near $x = 0$, $y = 0$, where $f(0, 0) = 0$. Find the Taylor polynomial of degree 2 for $f(x, y)$ in powers of x and y.

* 16. Use series methods to find the value of the partial derivative $f_{112}(0, 0)$ given that $f(x, y) = \arctan(x + y)$.

* 17. Use series methods to evaluate

$$\frac{\partial^{4n}}{\partial x^{2n}\partial y^{2n}} \frac{1}{1 + x^2 + y^2}\bigg|_{(0,0)}.$$

Chapter Review

Key Ideas

- **What do the following sentences and phrases mean?**
  - $\mathcal{S}$ is the graph of $f(x, y)$.
  - $\mathcal{C}$ is a level curve of $f(x, y)$.
  - $\lim_{(x,y)\to(a,b)} f(x, y) = L$.
  - $f(x, y)$ is continuous at (a, b).
  - the partial derivative $(\partial/\partial x) f(x, y)$
  - the tangent plane to $z = f(x, y)$ at (a, b)
  - pure second partials
  - mixed second partials
  - $f(x, y)$ is a harmonic function.
  - $L(x, y)$ is the linearization of $f(x, y)$ at (a, b).
  - the differential of $z = f(x, y)$
  - $f(x, y)$ is differentiable at (a, b).
  - the gradient of $f(x, y)$ at (a, b)
  - the directional derivative of $f(x, y)$ at (a, b) in direction $\mathbf{v}$
  - the Jacobian determinant $\partial(x, y)/\partial(u, v)$
- **Under what conditions are two mixed partial derivatives equal?**
- **State the Chain Rule for $z = f(x, y)$, where $x = g(u, v)$, and $y = h(u, v)$.**
- **Describe the process of calculating partial derivatives of implicitly defined functions.**
- **What is the Taylor series of $f(x, y)$ about (a, b)?**

Review Exercises

1. Sketch some level curves of the function $x + \dfrac{4y^2}{x}$.
2. Sketch some isotherms (curves of constant temperature) for the temperature function
$$T = \frac{140 + 30x^2 - 60x + 120y^2}{8 + x^2 - 2x + 4y^2} \quad (°\text{C}).$$
What is the coolest location?
3. Sketch some level curves of the polynomial function $f(x, y) = x^3 - 3xy^2$. Why do you think the graph of this function is called a *monkey saddle*?
4. Let $f(x, y) = \begin{cases} \dfrac{x^3}{x^2 + y^2}, & \text{if } (x, y) \neq (0, 0) \\ 0, & \text{if } (x, y) = (0, 0). \end{cases}$
Calculate each of the following partial derivatives or explain why it does not exist: $f_1(0, 0)$, $f_2(0, 0)$, $f_{21}(0, 0)$, $f_{12}(0, 0)$.

5. Let $f(x, y) = \dfrac{x^3 - y^3}{x^2 - y^2}$. Where is $f(x, y)$ continuous? To what additional set of points does $f(x, y)$ have a continuous extension? In particular, can f be extended to be continuous at the origin? Can f be defined at the origin in such a way that its first partial derivatives exist there?

6. The surface $\mathcal{S}$ is the graph of the function $z = f(x, y)$, where $f(x, y) = e^{x^2-2x-4y^2+5}$.

(a) Find an equation of the tangent plane to $\mathcal{S}$ at the point $(1, -1, 1)$.

(b) Sketch a representative sample of the level curves of the function $f(x, y)$.

7. Consider the surface $\mathcal{S}$ with equation $x^2 + y^2 + 4z^2 = 16$.

(a) Find an equation for the tangent plane to $\mathcal{S}$ at the point (a, b, c) on $\mathcal{S}$.

(b) For which points (a, b, c) on $\mathcal{S}$ does the tangent plane to $\mathcal{S}$ at (a, b, c) pass through the point $(0, 0, 4)$? Describe this set of points geometrically.

(c) For which points (a, b, c) on $\mathcal{S}$ is the tangent plane to $\mathcal{S}$ at (a, b, c) parallel to the plane $x + y + 2\sqrt{2}z = 97$?

8. Two variable resistors R_1 and R_2 are connected in parallel so that their combined resistance R is given by

$$\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2}.$$

If $R_1 = 100$ ohms $\pm 5\%$ and $R_2 = 25$ ohms $\pm 2\%$, by approximately what percentage can the calculated value of their combined resistance $R = 20$ ohms be in error?

9. You have measured two sides of a triangular field and the angle between them. The side measurements are 150 m and 200 m, each accurate to within ± 1 m. The angle measurement is $30°$, accurate to within $\pm 2°$. What area do you calculate for the field, and what is your estimate of the maximum percentage error in this area?

10. Suppose that $T(x, y, z) = x^3y + y^3z + z^3x$ gives the temperature at the point (x, y, z) in 3-space.

(a) Calculate the directional derivative of T at $(2, -1, 0)$ in the direction toward the point $(1, 1, 2)$.

(b) A fly is moving through space with constant speed 5. At time $t = 0$ the fly crosses the surface $2x^2 + 3y^2 + z^2 = 11$ at right angles at the point $(2, -1, 0)$, moving in the direction of increasing temperature. Find dT/dt at $t = 0$ as experienced by the fly.

11. Consider the function $f(x, y, z) = x^2y + yz + z^2$.

(a) Find the directional derivative of f at $(1, -1, 1)$ in the direction of the vector $\mathbf{i} + \mathbf{k}$.

(b) An ant is crawling on the plane $x + y + z = 1$ through $(1, -1, 1)$. Suppose it crawls so as to keep f constant. In what direction is it going as it passes through $(1, -1, 1)$?

(c) Another ant crawls on the plane $x + y + z = 1$, moving in the direction of the greatest rate of increase of f. Find its direction as it goes through $(1, -1, 1)$.

12. Let $f(x, y, z) = (x^2 + z^2)\sin\dfrac{\pi xy}{2} + yz^2$. Let P_0 be the point $(1, 1, -1)$.

(a) Find the gradient of f at P_0.

(b) Find the linearization $L(x, y, z)$ of f at P_0.

(c) Find an equation for the tangent plane at P_0 to the level surface of f through P_0.

(d) If a bird flies through P_0 with speed 5, heading directly toward the point $(2, -1, 1)$, what is the rate of change of f as seen by the bird as it passes through P_0?

(e) In what direction from P_0 should the bird fly at speed 5 to experience the greatest rate of increase of f?

13. Verify that for any constant k the function $u(x, y) = k\big(\ln\cos(x/k) - \ln\cos(y/k)\big)$ satisfies the *minimal surface equation*

$$(1 + u_x^2)u_{yy} - uu_xu_yu_{xy} + (1 + u_y^2)u_{xx} = 0.$$

14. The equations $F(x, y, z) = 0$ and $G(x, y, z) = 0$ can define any two of the variables x, y, and z as functions of the remaining variable. Show that

$$\frac{dx}{dy}\frac{dy}{dz}\frac{dz}{dx} = 1.$$

15. The equations $\begin{cases} x = u^3 - uv \\ y = 3uv + 2v^2 \end{cases}$ define u and v as functions of x and y near the point P where $(u, v, x, y) = (-1, 2, 1, 2)$.

(a) Find $\dfrac{\partial u}{\partial x}$ and $\dfrac{\partial u}{\partial y}$ at P.

(b) Find the approximate value of u when $x = 1.02$ and $y = 1.97$.

16. The equations $\begin{cases} u = x^2 + y^2 \\ v = x^2 - 2xy^2 \end{cases}$ define x and y implicitly as functions of u and v for values of (x, y) near $(1, 2)$ and values of (u, v) near $(5, -7)$.

(a) Find $\dfrac{\partial x}{\partial u}$ and $\dfrac{\partial y}{\partial u}$ at $(u, v) = (5, -7)$.

(b) If $z = \ln(y^2 - x^2)$, find $\dfrac{\partial z}{\partial u}$ at $(u, v) = (5, -7)$.

Challenging Problems

1. (a) If the graph of a function $f(x, y)$ that is differentiable at (a, b) contains part of a straight line through (a, b), show that the line lies in the tangent plane to $z = f(x, y)$ at (a, b).

(b) If $g(t)$ is a differentiable function of t, describe the surface $z = yg(x/y)$ and show that all its tangent planes pass through the origin.

2. A particle moves in 3-space in such a way that its direction of motion at any point is perpendicular to the level surface of

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

through that point. If the path of the particle passes through the point $(1, 1, 8)$, show that it also passes through $(2, 4, 1)$. Does it pass through $(3, 7, 0)$?

3. **(The Laplace operator in spherical coordinates)** If $u(x, y, z)$ has continuous second partial derivatives and

$$v(\rho, \phi, \theta) = u(\rho \sin\phi \cos\theta,\ \rho \sin\phi \sin\theta,\ \rho \cos\phi),$$

show that

$$\frac{\partial^2 v}{\partial \rho^2} + \frac{2}{\rho}\frac{\partial v}{\partial \rho} + \frac{\cot\phi}{\rho^2}\frac{\partial v}{\partial \phi} + \frac{1}{\rho^2}\frac{\partial^2 v}{\partial \phi^2} + \frac{1}{\rho^2 \sin^2\phi}\frac{\partial^2 v}{\partial \theta^2} = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2}.$$

You can do this by hand, but it is a lot easier using computer algebra.

4. **(Spherically expanding waves)** If f is a twice differentiable function of one variable and $\rho = \sqrt{x^2 + y^2 + z^2}$, show that $u(x, y, z, t) = \dfrac{f(\rho - ct)}{\rho}$ satisfies the three-dimensional wave equation

$$\frac{\partial^2 u}{\partial t^2} = c^2 \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} \right).$$

What is the geometric significance of this solution as a function of increasing time t? *Hint:* you may want to use the result of Exercise 3. In this case $v(\rho, \phi, \theta)$ is independent of ϕ and θ.