Chapter 16

Partial Derivatives

The use of contour lines to help understand a function whose domain is part of the plane goes back to the year 1774. A group of surveyors had collected a large number of the elevations of points on Mount Schiehalli in Scotland. They were doing this in order to estimate its mass and by its gravitational attraction, the mass of the earth. They asked the mathematician Charles Hutton for help in using the data entered as a map. Hutton saw that if he connected points on the map that showed the same elevation, the resulting curves — contour lines — suggested the shape of the mountain.

Reference: Bill Bryson, *A Short History of Nearly Everything*, Broadway Books, New York, 2003, p. 57.

16.1 Picturing a Function of Several Variables

The graph of $y = f(x)$, a function of just one variable, x, is a curve in the xy-plane. The graph of a function of two variables, $z = f(x, y)$ is a surface in space. It consists of the points (x, y, z) for which $z = f(x, y)$. For instance, if $z = 2x + 3y$, the graph is the plane $2x + 3y - z = 0$.

A vector field in the xy-plane is a vector-valued function of x and y. We pictured it by drawing a few vectors with their tails placed at the arguments.

This section describes some of the ways of picturing a scalar-valued functions of two or three variables.

Contour Lines

This is similar to what we did for vector fields.

For a function, $z = f(x, y)$, the simplest method is to attach at some point (x, y) the value of the function, $z = f(x, y)$. For instance, if $z = xy$, Figure 16.1.1 shows this method. This conveys a sense of the function. Its

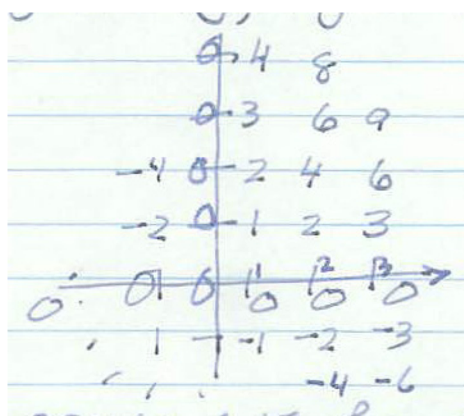


Figure 16.1.1:

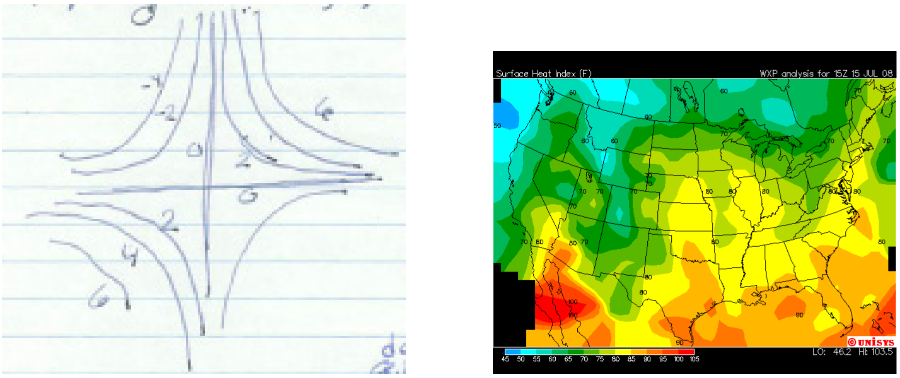
values are positive in the first and third quadrants, negative in the second and fourth. For (x, y) far from the origin near the lines $y = x$ or $y = -x$ the values are large.

Rather than attach the values at points, we could indicate all the points where the function has a specific fixed value. In other words we could graph, for a constant c, all the points where $f(x, y) = c$. Such as graph is called a **contour** or **level curve**.

contours and level curves

For the function $z = xy$, the contours are hyperbolas $xy = c$. In Figure 16.1.2(a) the contours corresponding to $c = 2,\ 4,\ 6,\ 0,\ -2,\ -4,\ -6$ are shown.

Many newspapers publish a daily map showing the temperature throughout the nation with the aid of contour lines. Figure 16.1.2(b) is an example.



(a) (b)

Figure 16.1.2:

At a glance you can see where it is hot or cold and in what direction to travel to warm up or cool off.

Traces

SHERMAN: xrcs Katrina wind / pressure?

Another way to get some idea of what the surface $z = f(x, y)$ looks like is to sketch the intersection of various planes with the surface. These intersections (or cross sections) are called **traces**.

For instance, Figure 16.1.3 exhibits the notion of a trace by a plane parallel to the xy-coordinate plane, namely, the plane $z = k$ This trace is an exact copy of the contour $f(x, y) = k$, as shown in Figure 16.1.3.

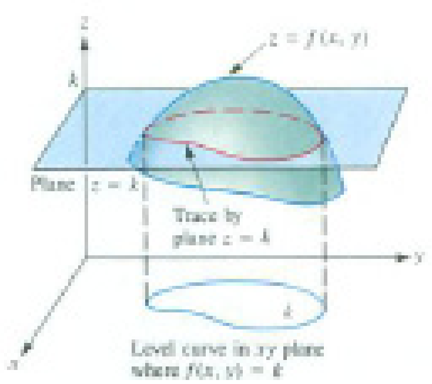


Figure 16.1.3:

Doug: maybe $z = x^2 - y^2$ is better? SHERMAN: I'm OK with xy, it's just a matter of perspective.

EXAMPLE 1 Sketch the traces of the surface $z = xy$ with the planes

1. $z = 1$,

2. $x = 1$,
3. $y = x$,
4. $y = -x$,
5. $x = 0$.

SOLUTION

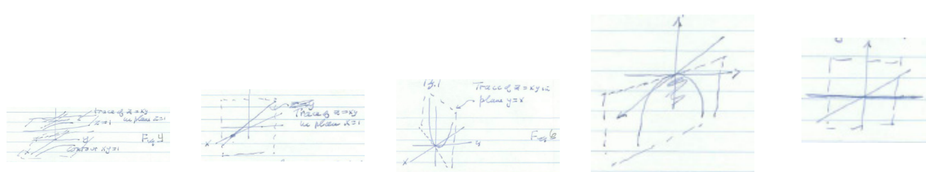
1. The trace with the plane $z = 1$ is shown in Figure 16.1.4. For points (x, y, z) on this trace $xy = 1$. The trace is a hyperbola. In fact, it is just the contour line $xy = 1$ in the xy plane raised by one unit as in Figure 16.1.4(a)

2. The trace in the plane $x = 1$ satisfies the equation $z = 1 \cdot y = y$. It is a straight line, shown in Figure 16.1.4(b)

3. The trace in the plane $y = x$ satisfies the equation $z = x^2$. It is the parabola shown in Figure 16.1.4(c).

4. The trace in the plane $y = -x$ satisfies the equation $z = x(-x) = -x^2$. It is an "upside-down" parabola, shown in Figure 16.1.4(d).

5. The intersection with the coordinate plane $x = 0$ satisfies the equation $z = 0 \cdot y = 0$. It is the y-axis, shown in Figure 16.1.4(e).

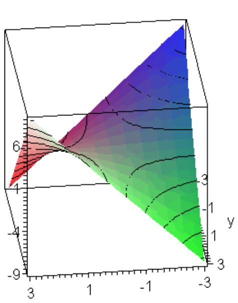


(a) (b) (c) (d) (e)

Figure 16.1.4:

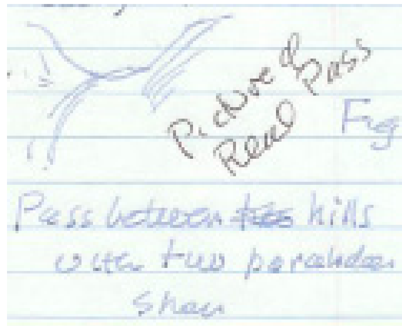
So the surface can be viewed as made up of lines, or of parabolas or of hyperbolas.

The surface $z = xy$ is shown in Figure 16.1.5 with some of the traces drawn on it. ◇



The surface $z = xy$ looks like a saddle or the pass between two hills, as shown in Figure 16.1.6.

(a)



(b)

Figure 16.1.6:

Functions of Three Variables

The graph of $y = f(x)$ consists of certain points in the xy plane. The graph of $z = f(x, y)$ consists of certain points in the xyz space. But what if we have a function of three variables, $u = f(x, y, x)$? (The volume V of a box of sides x, y, z is given by the equation $V = xyz$; this is an example of the function of three variables.) We cannot graph the set of points (x, y, z, u) where $u = f(x, y, z, u)$ since we live in space of only three dimensions. What we could do is pick a constant k and draw the "level surfaces," the set of points where $f(x, y, z) = k$. Varying k may give an idea of this function's behavior, just as varying the k of $f(x, y) = k$ yields information about the behavior of a function of two variables.

For example, let $T = f(x, y, z)$ be the temperature (Fahrenheit) at the point (x, y, z). Then the level surface

$$68 = f(x, y, z)$$

consists of all points where the temperature is 68°.

EXAMPLE 2 Describe the level surfaces of the function $u = x^2 + y^2 + z^2$.
SOLUTION For each k we examine the equation $u = x^2 + y^2 + z^2$. If k is negative, there are no points in the "level surface." If $k = 0$, there is only one point, the origin $(0, 0, 0)$. If $k = 1$, the equation $1 = x^2 + y^2 + z^2$, which describes a sphere of radius 1 center $(0, 0, 0)$. If k is positive, the level surface $f(x, y, z) = k$ is a sphere of radius $\sqrt{k}$, center $(0, 0, 0)$. See Figure 16.1.7 ⋄

Summary

We introduced the idea of a function of two variables $z = f(P)$ is in some region in the xy plane. The graph of $z = f(P)$ is usually a surface. But it is

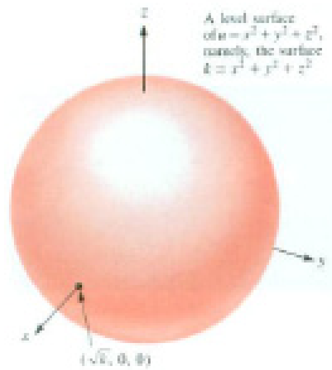


Figure 16.1.7:

often more useful to sketch a few of its level curves than to sketch that surface. Each level curve is the projection of a trace of the surface in a plane of the form $z = k$. Note that at all points (x, y) on a level curve the function have the same value. In other words, the function f is constant on a level curve.

In particular, we used level curves to analyze the function $z = xy$ whose graph is a saddle.

For functions of three variables $u = (x, y, z)$, we defined level surfaces. When considered on a level surface, $k = f(x, y, z)$ such a function is constant, with value k.

EXERCISES for Section 16.1 *Key:* R–routine, M–moderate, C–challenging

In Exercises 1 to 10, graph the given function.

1.[R] $f(x, y) = y$

2.[R] $f(x, y) = x + 1$

3.[R] $f(x, y) = 3$

4.[R] $f(x, y) = -2$

5.[R] $f(x, y) = x^2$

6.[R] $f(x, y) = y^2$

7.[R] $f(x, y) = x + y + 1$

8.[R] $f(x, y) = 2x - y + 1$

9.[R] $f(x, y) = x^2 + 2y^2$

10.[M] $f(x, y) = \sqrt{x^2 + y^2}$

In Exercises 11 to 14 draw for the given functions the level curves corresponding to the values −1, 0, 1, and 2 (if they are not empty).

11.[R] $f(x, y) = x + y$

12.[R] $f(x, y) = x + 2y$

13.[R] $f(x, y) = x^2 + 2y^2$

14.[R] $f(x, y) = x^2 - 2y^2$

In Exercises 15 to 18 draw the level curves for the given functions that pass through the given points.

15.[R] $f(x, y) = x^2 + y^2$ through $(1, 1)$ HINT: First compute $f(1, 1)$.

16.[R] $f(x, y) = x^2 + 3y^2$ through $(1, 2)$

17.[R] $f(x, y) = x^2 - y^2$ through $(3, 2)$

18.[R] $f(x, y) = x^2 - y^2$ through $(2, 3)$

19.[R]

(a) Draw the level curves for the functions $f(x, y) = x^2 + y^2$ corresponding to the values $k = 0, 1, \ldots, 9$.

(b) By inspection of the curves in (a), decide where the functions changing most rapidly. Explain why you think so.

20.[R] Let $f(P)$ be the average daily solar radiation at the point P (measured in langleys). The level curves corresponding to 350, 400, 450, and 500 langley are shown in Figure 16.1.8.

Figure 16.1.8:

(a) What can be said about the ratio between the maximum and minimum solar radiation at points in the United States?

(b) Why are there rather sharp bends in the level curves in two areas?

21.[R] Let $u = g(x, y, z)$ be a function of three variables. Describe the level surface $g(x, y, z) = 1$ if $g(x, y, z)$ is

(a) $x + y + z$

(b) $x^2 + y^2 + z^2$

(c) $x^2 + y^2 - z^2$

(d) $x^2 - y^2 - z^2$ HINT: For (c) and (d) are examples of quadric surfaces.

22.[R]

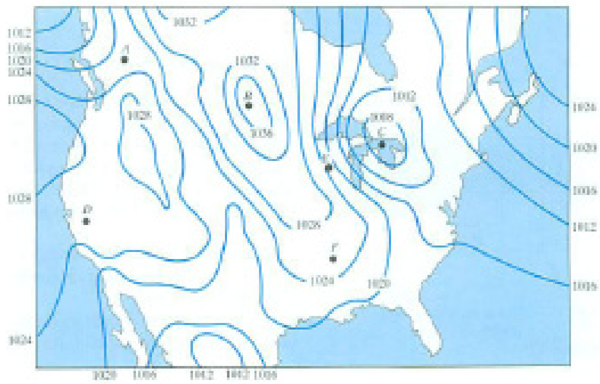


Figure 16.1.9:

The daily weather map shows the barometric pressure function by a few well-chosen level curves (called *isobars*), as in Figure 16.1.9. In this case, the function is 'pressure at (x, y)."

(a) Where is the lowest pressure?

(b) Where is the highest pressure?

(c) Where do you think the wind at ground level is the fastest? Why?

23.[R] A map of August, 26, 2005 showing isobars and wind vectors, day of Katrina and some questions.

24.[R] Questions about the map in Figure 16.1.2(b).

25.[M]

(a) Sketch the surface $z = x^2 +$

(b) Show that all the traces by p xz plane are parabolas.

(c) Show that the parabolas in (l (So the surface is made up c las.)

(d) What kind of curve is a trac to the xy plane?

26.[M] Consider the surface $z = x$ of curve is produced by a trace by

(a) the xy plane,

(b) the xz palne,

(c) the yz plane.

27.[C]

(a) Is the parabola $y = x^2$ congru $y = 4x^2$?

(b) Is the parabola $y = x^2$ simi $y = 4x^2$? (One figure is simil is simply the other magnified in all direction.)

16.2 The Many Derivatives of $f(x, y)$.

The notions of limit, continuity and derivative carry over with similar definitions from functions $f(x)$ of one variable to functions of several variables $f(x, y)$. However, the derivatives of functions of several variable involves some new ideas.

SHERMAN: There are more new ideas for limits than derivatives. In fact, partial

Limits and Continuity of $f(x, y)$

The **domain** of function $f(x, y)$ is the set of points where it is defined. The domain of $f(x, y) = x + y$ is the entire xy plane. The domain of $f(x, y) = \sqrt{1 - x^2 - y^2}$ is much smaller. In order for the square root of $1 - x^2 - y^2$ to be defined, $1 - x^2 - y^2$ must not be negative. In other words, we must have $x^2 + y^2 \leq 1$. The domain is the disk bounded by the circle $x^2 + y^2 = 1$, shown in Figure 16.2.1.

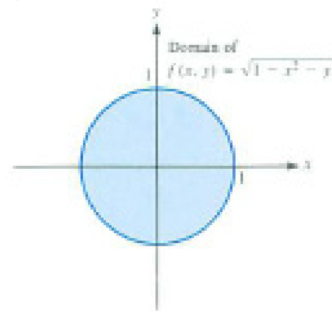


Figure 16.2.1:

A point P_0 is on the **boundary** of a set if every disk centered at P_0, no matter how small, contains points in the set and points not in the set. (See Figure 16.2.3.) The boundary of the circle $x^2+y^2 \leq 1$ is the circle $x^2+y^2 = 1$. The domain of $f(x, y) = \sqrt{1 - x^2 - y^2}$ includes every point on its boundary.

The domain of $f(x, y) = 1/\sqrt{1 - x^2 - y^2}$ is even smaller. Now we must not let $1-x^2-y^2$ be 0 or negative. The domain of $1/\sqrt{1 - x^2 - y^2}$ consists of the points (x, y) such that $x^2 + y^2 < 1$. It is the disk in Figure 16.2.1 *without* its boundary.

The function $f(x, y) = 1/(y - x)$ is defined everywhere except on the line $y - x = 0$. Its domain is the xy plane from which the line $y = x$ is removed. (See Figure 16.2.2.)

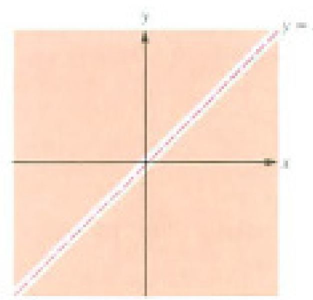
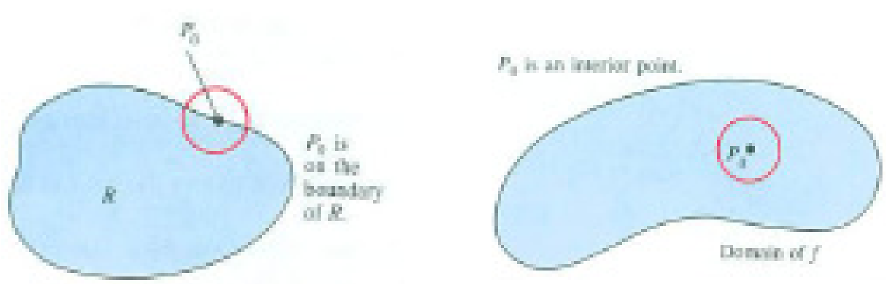


Figure 16.2.2:

The domain of a function of interest to us will either be the entire xy plane or some region bordered by curves or lines, or perhaps such a region with a few points omitted. Let P_0 be a point in the domain of a function f. If there is



(a) (b)

Figure 16.2.3:

a disk with center P_0 that lies within the domain of f, we call P_0 an **interior** point of the domain. (See Figure 16.2.3(b).) When P_0 is an interior point of the domain of f, we know that $f(P)$ is defined for all points P sufficiently near P_0. Every point P_0 in the domain not on its boundary is an interior point. A set R is called **open** if each point P of R is an interior point of R. The entire xy plane is open. So is any disk without its circumference. More generally, the set of points inside some closed curve but not on it forms an open set.

The definition of the limit of $f(x, y)$ as (x, y) approaches $P_0 = (a, b)$ will not come as a surprise.

DEFINITION *(Limit of $f(x, y)$ at $P_0 = (a, b)$)* Let f be a function defined at least at every point in some disk with center P_0, except perhaps at P_0. If there is a number L such that $f(P)$ approaches L whenever P approaches P_0 we call L the **limit of $f(P)$ as P approaches P_0**. We write

$$\lim_{P \to P_0} f(P) = L$$

or

$$f(P) \to L \qquad \text{as} \qquad P \to P_0.$$

We also write

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

For most of the functions of interest the limit will always exist throughout its domain. However, even a formula that is easily defined may not have a limit at some points.

EXAMPLE 1 Let $f(x, y) = \frac{x^2 - y^2}{x^2 + y^2}$. Determine whether $\lim_{P \to (0,0)} f(P)$ exists.

SOLUTION The function is not defined at $(0, 0)$. When (x, y) is near $(0, 0)$, both the numerator and denominator of $(x^2 - y^2)/(x^2 + y^2)$ are small numbers. There are, as in Chapter 2, two influences. The numerator is pushing the quotient towards 0 while the denominator is influencing the quotient to be large. We must be careful.

We try a few inputs near $(0, 0)$. For instance, $(0.01, 0)$ is near $(0, 0)$ and

$$f(0.01, 0) = \frac{(0.01)^2 - 0^2}{(0.01)^2} + 0^2 = 1$$

Also, $(0, 0.01)$ is near $(0, 0)$ and

$$f(0, 0.01) = \frac{0^2 - (0.01)^2}{0^2 + (0.01)^2} = -1$$

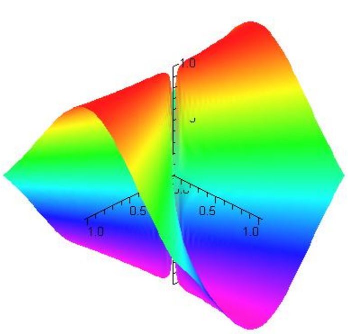


Figure 16.2.4:

More generally, for $x \neq 0$,

$$f(x, 0) = 1;$$

while, for $y \neq 0$,

$$f(0, y) = -1$$

Since x can be as near 0 as we please and y can be as near 0 as we please, it is *not* the case that $\lim_{P \to (0,0)} f(P)$ exists. Figure 16.2.4 shows the graph of $z = \frac{x^2-y^2}{x^2+y^2}$. ◇

Continuity of $f(x, y)$ at $P_0 = (a, b)$

With only slight changes, the definition of continuity for $f(x)$ in Section 2.4 easily generalizes to the definition of continuity for $f(x, y)$.

> **DEFINITION** *(Continuity of $f(x, y)$ at $P_0 = (a, b)$).* Assume that $f(P)$ is defined throughout some disk with center P_0. Then f is **continuous** at P_0 if $\lim_{P \to P_0} f(P) = f(P_0)$.
>
> This means
>
> 1. $f(P_0)$ is defined (that is, P_0 is in the domain of f),
> 2. $\lim_{P \to P_0} f(P)$ exists, and
> 3. $\lim_{P \to P_0} f(P) = f(P_0)$.
>
> Continuity at a point on the boundary of the domain can be defined similarly. A function $f(P)$ is **continuous** if it is continuous at every point in its domain.

EXAMPLE 2 Determine whether $f(x, y) = \frac{x^2-y^2}{x^2+y^2}$ is continuous at $(1, 1)$.
SOLUTION This is the function explored in Example 1. First, $f(1, 1)$ is defined. (It equals 0.) Second, $\lim_{(x,y) \to (1,1)} \frac{x^2-y^2}{x^2+y^2}$. (It is $\frac{0}{2} = 0$.) Third, $\lim_{(x,y) \to (1,1)} f(x, y) = f(1, 1)$.

Hence, $f(x, y)$ is continuous at $(1, 1)$. ◇

In fact, the function of Example 2 is continuous at every point (x, y) in its domain. We do not need to worry about the behavior of $f(x, y)$ when (x, y) is near $(0, 0)$ because $(0, 0)$ is *not* in the domain. Since $f(x, y)$ is continuous at every point *in its domain*, it is a continuous function.

The Two Partial Derivatives of $f(x, y)$

Let (a, b) be a point on the domain of $f(x, y)$. The trace on the surface $z = f(x, y)$ by a plane through (a, b) and parallel to the z-axis is a curve, as shown in Figure 16.2.5.

If f is well behaved at the point $P = (a, b, f(a, b))$ the trace has a slope. This slope depends on the plane through (a, b). In this section we consider only the two planes parallel to the coordinate planes $y = 0$ and $x = 0$. In the next section we treat the general cases.

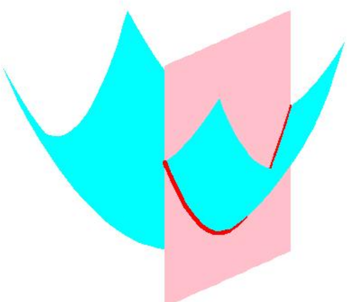


Figure 16.2.5:

Consider the function $f(x, y) = x^2y^3$. If we hold y constant and differentiate with respect to x, we obtain $d(x^2y^3)/dx = 2xy^3$. This derivative is called the "partial derivative" of x^2y^3 with respect to x. We could hold x fixed instead and find the derivative of x^2y^3 with respect to y, that is, $d(x^2y^3)/dy = 3x^2y^2$. This derivative is called the "partial derivative" of x^2y^3 with respect to y. This example introduces the general idea of partial derivative. First we define them. Then we will see what they mean in terms of slope and rate of change.

DEFINITION *(Partial derivatives.)* Assume that the domain of $f(x, y)$ includes the region within some disk with center (a, b). If

$$\lim_{\Delta x \to 0} \frac{f(a + \Delta x, b) - f(a, b)}{\Delta x}$$

exists, this limit is called the **partial derivative of f with respect to x** at (a, b). Similarly, if

$$\lim_{\Delta x \to 0} \frac{f(a, b + \Delta y) - f(a, b)}{\Delta y}$$

exists, it is called the **partial derivative of f with respect to y** at $9a, b)$.

Notations for partial derivatives.

The following notations are used for the partial derivatives of $z = f(x, y)$ with respect to x:

$$\frac{\partial z}{\partial x}, \frac{\partial f}{\partial y}, f_x, f_1, \quad \text{or} \quad z_x.$$

And the following are used for partial derivative of $z = f(x, y)$ with respect to y:

$$\frac{\partial z}{\partial y}, \frac{\partial f}{\partial y}, f_y, f_2, \quad \text{or} \quad z_y.$$

Since physicists and engineers use the subscript notation in study of vectors, they prefer to use

$$\frac{\partial f}{\partial x} \quad \text{and} \quad \frac{\partial f}{\partial y}$$

to denote the two partial derivatives. The symbol $\partial\partial f/\partial x$ may be viewed as "the rate at which the function $f(x,y)$ changes when x varies and y is kept fixed." The symbol $\partial f/\partial y$ records "the rate at which the function $f(x,y)$ changes when y varies and x is kept fixed."

The value of $\partial f/\partial x$ at (a,b) is denoted

$$\frac{\partial f}{\partial x}(a,b) \quad \text{or} \quad \left.\frac{\partial f}{\partial x}\right|_{(a,b)}.$$

In the middle of a sentence, we will write it as $f_x(a,b)$ or $\partial f/\partial x(a,b)$.

EXAMPLE 3 If $f(x,y) = \sin(x^2y)$, find

1. $\partial f/\partial x$,
2. $\partial f/\partial y$, and
3. $\partial f/\partial y$ at $(1,\pi/4)$.

SOLUTION

1. To find $\frac{\partial}{\partial x}(\sin x^2y)$, differentiate with respect to x, keeping y constant:

$$\begin{aligned} \tfrac{\partial}{\partial x}(\sin x^2y) &= \cos(x^2y)\tfrac{\partial}{\partial x}(x^2y) && \text{chain rule} \\ &= \cos(x^2y)(2xy) && y \text{ is constant} \\ &= 2xy\cos(x^2y). \end{aligned}$$

2. To find $\frac{\partial}{\partial y}(\sin x^2y)$, differentiate with respect to y, keeping x constant:

$$\begin{aligned} \tfrac{\partial}{\partial y}(\sin x^2y) &= \cos(x^2y)\tfrac{\partial}{\partial y}(x^2y) && \text{chain rule} \\ &= \cos(x^2y)(x^2) && x \text{ is constant} \\ &= x^2\cos(x^2y). \end{aligned}$$

3. By (b)

$$\tfrac{\partial f}{\partial y}(1,\pi/4) = \left. x^2\cos(x^2y)\right|_{(1,\pi/4)} = 1^2\cos\left(1^2\tfrac{\pi}{4}\right) = \tfrac{\sqrt{2}}{2}.$$

◇

As Example 3 shows, since partial derivatives are really ordinary derivatives, the procedures for computing derivatives of a function $f(x)$ of a single variable carry over to functions of two variables.

Higher-Order Partial Derivatives

Just as there are derivatives of derivatives, so are there partial derivatives of partial derivatives. For instance, if

$$z = 2x + 5x^4y^7,$$

then

$$\frac{\partial z}{\partial x} = 2 + 20x^3y^7 \qquad \text{and} \qquad \frac{\partial z}{\partial y} = 35x^4y^6.$$

We may go on and compute the partial derivatives of $\partial z/\partial x$ and $\partial z/\partial y$:

$$\begin{array}{rclrcl} \frac{\partial}{\partial x}\left(\frac{\partial z}{\partial x}\right) & = & 60x^2y^7 & \frac{\partial}{\partial y}\left(\frac{\partial z}{\partial y}\right) & = & 140x^3y^6 \\ \frac{\partial}{\partial x}\left(\frac{\partial z}{\partial y}\right) & = & 140x^3y^6 & \frac{\partial}{\partial y}\left(\frac{\partial z}{\partial y}\right) & = & 210x^4y^5. \end{array}$$

There are four partial derivatives of the second order:

$$\frac{\partial}{\partial x}\left(\frac{\partial z}{\partial x}\right), \frac{\partial}{\partial y}\left(\frac{\partial z}{\partial x}\right), \frac{\partial}{\partial y}\left(\frac{\partial z}{\partial y}\right), \frac{\partial}{\partial x}\left(\frac{\partial z}{\partial y}\right).$$

These are usually denoted, in the same order, as

$$\frac{\partial^2 z}{\partial x^2}, \frac{\partial^2 z}{\partial y \partial x}, \frac{\partial^2 z}{\partial y^2}, \frac{\partial^2 z}{\partial x \partial y}.$$

To compute $\partial^2 z/\partial x \partial y$, you first differentiate with respect to y, then with respect to x. To compute $\partial^2 z/\partial y \partial x$, you first differentiate with respect to x, then with respect to y. In both cases, "differentiate from right to left in the order that the variables occur."

The partial derivative $\frac{\partial f}{\partial x}$ is also denoted f_x and $\frac{\partial f}{\partial y}$ is denoted f_y. The second partial derivative $\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial (f_y)}{\partial x} = (f_y)_x$ is denoted f_{yx}. In this case you differentiate from left to right, "first f_y, then $(f_y)_x$." In short, $f_{yx} = (f_y)_x$, $f_{yy} = (f_y)_y$, and $f_{xy} = (f_x)_y$. In both notations the mixed partial is computed in the order that resembles its definition (with the parentheses removed), Thus

The subscript notation, f_{yx}, is generally preferred in the midst of other text.

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) \qquad \text{and} \qquad f_{xy} = (f_x)_y$$

are the two different mixed second partial derivatives of f.

Equality of the mixed partials

In the computations just done, the two mixed partials z_{xy} and z_{yx} are equal. For the functions commonly encountered, the two mixed partials are equal. (For a proof, see Appendix K.)

SHERMAN: V had an appedix on interchanging limits. How will we deal with this in VI?

Exercise 27 presents a function for which the two mixed particles are not equal. Such a special case mathematicians call "pathological", though the function does not view itself as sick.

EXAMPLE 4 Compute $\frac{\partial^2 z}{\partial x^2} = f_{xx}$, $\frac{\partial^2 z}{\partial y \partial x} = f_{xy}$, and $\frac{\partial^2 z}{\partial x \partial y} = z_{yx}$ for $z = y\cos(xy)$.

SOLUTION First compute

$$\frac{\partial^2 z}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\partial z}{\partial x}\right) = \frac{\partial}{\partial x}(-y^2\sin(xy)) = -y^3\cos(xy).$$

Then

$$\frac{\partial^2 z}{\partial y \partial x} = \frac{\partial}{\partial y}\left(\frac{\partial z}{\partial x}\right) = \frac{\partial}{\partial y}(-y^2\cos(xy)) = -2y\sin(xy) - xy^2\cos(xy).$$

Finally,

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

Notice that while the work required to compute the mixed partials is very different, the two derivatives are, as expected, are equal. ◇

Functions of More Than Two Variables

A quantity may depend on more than two variables. For instance, the volume of a box depends on three variables: the length l, width w, and height h, $V = lwh$. The "chill factor" depends on the temperature, humidity, and wind velocity. The temperature T at any point in the atmosphere is a function of the three space coordinates, x, y, and z: $T = f(x, y, z)$.

To differentiate, hold all variables constant except one.

The notions and notations of partial derivatives carry over to functions of more than two variables. If $u = f(x, y, z, t)$, there are four first-order partial derivatives. For instance, the partial derivative of u with respect to x, holding y, z, and t fixed, is denoted

$$\frac{\partial u}{\partial x}, \ \frac{\partial f}{\partial x}, \ u_x, \quad \text{etc.}$$

Higher-ordered partial derivatives are defined and denoted similarly. Many basic problems in chemistry and physics, such as vibrating string are examined in terms of equations involving partial derivatives (known as PDEs).

Insert CIE on the Vibrating String.

Summary

We defined limit, continuity, and derivatives for functions of several variables. These notions are all closely related to the one variable versions.

A key difference is that a partial derivative with respect to one variable, say x, is found by treating all other variables as constants and applying the standard differentiation rules with respect to x. Higher-order partial derivatives are also defined much like higher-order derivatives. An important property of higher-order partial derivatives is that the order in which the partial derivatives are applied can be important, but not for the functions usually met in applications.

EXERCISES for Section 16.2 *Key:* R–routine, M–moderate, C–challenging

In Exercises 1 to 8 evaluate the limits, if they exist.

1.[R] $\lim_{(x,y)\to(2,3)} \frac{x+y}{x^2+y^2}$

2.[R] $\lim_{(x,y)\to(1,1)} \frac{x^2}{x^2+y^2}$

3.[R] $\lim_{(x,y)\to(0,0)} \frac{x^2}{x^2+y^2}$

4.[R] $\lim_{(x,y)\to(0,0)} \frac{xy}{x^2+y^2}$

5.[R] $\lim_{(x,y)\to(2,3)} x^y$

6.[R] $\lim_{(x,y)\to(0,0)} (x^2)^y$

7.[R] $\lim_{(x,y)\to(0,0)} (1+xy)^{1/(xy)}$

8.[R] $\lim_{(x,y)\to(0,0)} (1+x)^{1/y}$

In Exercises 9 to 14, (a) describe the domain of the given functions and (b) state whether the functions are continuous.

9.[R] $f(x,y) = 1/(x+y)$

10.[R] $f(x,y) = 1/(x^2+2y^2)$

11.[R] $f(x,y) = 1/(9-x^2-y^2)$

12.[R] $f(x,y) =$ $\sqrt{x^2+y^2-25}$

13.[R] $f(x,y) = \sqrt{16-x^2-y^2}$

14.[R] $f(x,y) = \sqrt{49-x^2-y^2}$

In Exercises 15 to 20, find the boundary of the given region R.

15.[R] R consists of all points (x,y) such that $x^2+y^2 \le 1$.

16.[R] R consists of all points (x,y) such that $x^2+y^2 < 1$.

17.[R] R consists of all points (x,y) such that $1/(x^2+y^2)$ is defined.

18.[R] R consists of all points (x,y) such that $1/(x+y)$ is defined.

19.[R] R consists of all points (x,y) such that $y < x^2$.

20.[R] R consists of all points (x,y) such that $y \le x$.

In Exercises 21 to 24 concern the precise definition of $\lim_{(x,y)\to P_0} f(x,y)$.

21.[R] Let $f(x,y) = x+y$.

(a) Show that if $P = (x,y)$ lies within a distance 0.01 of $(1,2)$, then $|x-1| < 0.01$ and $|y-2| < 0.01$. (See Figure 16.2.6).

SHERMAN: Move some of these to Chapter Summary. Emphasis is on partial derivatives.

(b) Show that if $|x-1| < 0.01$ and $|y-2| < 0.01$, then $|f(x,y)-3| < 0.02$.

(c) Find a number $\delta > 0$ such that if $P = (x,y)$ is in the disk of center $(1,2)$ and radius δ, then $|f(x,y)-3| < 0.001$.

(d) Show that for any positive number ϵ, no matter how small, there is a positive number δ such that when $P = (x,y)$ is in the disk of radius δ and center $(1,2)$, then $|f(x,y)-3| < \epsilon$. (Give δ as a function of ϵ.)

(e) What may we conclude on the basis of (d)?

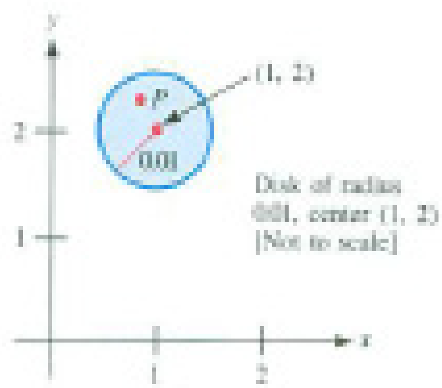


Figure 16.2.6:

22.[R] Let $f(x,y) = 2x+3y$.

(a) Find a disk with center $(1,1)$ such that whenever P is in that disk, $|f(P)-5| < 0.01$

(b) Let ϵ be any positive number. Show that there is a disk with center $(1,1)$ such that whenever P us in that disk, $|f(P)-5| < \epsilon$. (Give δ as a function of ϵ.)

(c) What may we conclude on the basis of (b)?

23.[R] Let $f(x,y) = s^2y/(x^4+2y^2)$.

(a) What is the domain of f?

(b) Fill in this table:

(x,y)	$(0.01, 0.01)$	$(0.01, 0.02)$	$(0.001, 0.003)$
$f(x,y)$			

(c) On the basis of (b), do you think $\lim_{P\to(0,0)} f(P)$ exists? If so, what is its value?

(d) Fill in this table:

(x,y)	$(0.5, 0.25)$	$(0.1, 0.01)$	$(0.001, 0.000001)$
$f(x,y)$			

(e) On the basis of (d), do you think $\lim_{P\to(0,0)} f(P)$ exists? If so, what is its value?

(f) Does $\lim_{P\to(0,0)} f(P)$ exist? If so, what is it? Explain.

24.[R] Let $f(x,y) = 5x^2y/(2x^4+3y^2)$.

(a) What is the domain of f?

(b) As P approaches $(0,0)$ on the line $y = 2x$, what happens to $f(P)$?

(c) As P approaches $(0,0)$ on the line $y = 3x$, what happens to $f(P)$?

(d) As P approaches $(0,0)$ on the parabola $y = x^2$, what happens to $f(P)$?

(e) Does $\lim_{P\to(0,0)} f(P)$ exist? If so, what is it? Explain.

25.[R] Show that for any polyn
equals P_{xy}. Suggestion: It is enoug
arbitrary monomial ax^my^n, where
and n are non-negative integers. T
n is 0 should be treated separately

26.[M] Let $T(x,y,z) = 1/\sqrt{x^2+$
is not the origin $(0,0,0)$. Show tha

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

This equation arises in the theory
show in Section 16.4.

27.[C] This exercise presents a fu
that its two mixed partial derivati
equal.

(a) Let $g(x,y) = \frac{x^2-y^2}{x^2+y^2}$ for
Show that $\lim_{k\to0}(\lim_{h\to0} g$
$\lim_{h\to0}(\lim_{k\to0} g(h,k)) = 1$.

(b) Let $f(x,y) = xyg(x,y)$ if $(x$
$f(0,0) = 0$. Show that $f(x,y$

(c) Show that $f_{xy}(0,0) = \lim_{k\to}$

(d) Show that $f_{xy}(0,0) = \lim_{k\to0}$

(e) Show that $f_{xy}(0,0) = -1$.

(f) Similarly, show that $f_{xy}(0,0$

(g) Show that in polar coordinat
the point (r,θ) is $r^2\sin(4\theta)/$

16.3 Change and the Chain Rule

For a function of one variable, $f(x)$, the change in the value of the function as the input changes from a to $a+\Delta x$ is approximately $f'(a)\Delta x$. In this section we estimate the change in $f(x,y)$ as (x,y) moves from (a,b) to $(a+\Delta x, b+\Delta y)$.

That type of estimate is the key to obtaining the chain rule for functions of several variables. We will find that the chain rule involves the *sum* of terms that resemble the product $\frac{dy}{du} \cdot \frac{du}{dx}$ that appear in the chain rule for a function of one variable.

Estimating the Change of Δf

Let $z = f(x,y)$ be a function of two variables with continuous partial derivatives at least throughout a disk centered at the point (a,b). We will express $\Delta f = f(a+\Delta x, b+\Delta y) - f(a,b)$ in terms of f_x and f_y. This change is shown in Figure 16.3.1. We can view this change as obtained in two steps. First, the

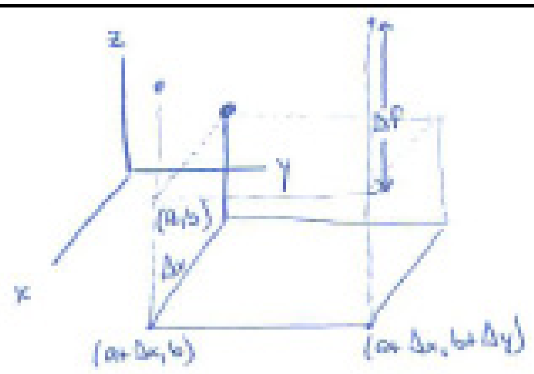


Figure 16.3.1:

change as x goes from a to $a+\Delta x$, that is, $f(a+\Delta x, b) - f(a,b)$. Second, the change from $f(a+\Delta x, b)$ to $f(a+\Delta x, b+\Delta y)$, as y changes from b to $b+\Delta y$.

In short,

$$\Delta f = (f(a+\Delta x, b) - f(a,b)) + (f(a+\Delta x, b+\Delta y) - f(a+\Delta, b)). \quad (16.3.1)$$

By the mean-value theorem, there is a number c_1 between a and $a+\Delta x$ such that

$$f(a+\Delta x, b) - f(a,b) = \frac{\partial f}{\partial x}(c_1, b)\Delta x \quad (16.3.2)$$

(16.3.1) is clear algebraically because the two $f(a+\Delta x, b)$ terms cancel.

Similarly, applying the mean-value theorem to the second bracket expression as (16.3.2), we see that there is a number c_3 between b and $b+\Delta y$ such that

$$f(a+\Delta x, b+\Delta y) - f(a+\Delta x, b) = \frac{\partial f}{\partial y}(a+\Delta x, c_2)\Delta y. \qquad (16.3.3)$$

Combining (16.3.1), (16.3.2) and (16.3.3) we obtain

$$\Delta f = \frac{\partial f}{\partial x}(c_1, b)\Delta x + \frac{\partial f}{\partial y}(z+\Delta x, c_2)\Delta y. \qquad (16.3.4)$$

When both Δx and Δy are small, the points (c_1, b) and $(a+\Delta x, c_2)$ are near the point (a, b). If we assume that the partial derivatives f_x are continuous at (a, b), then we may conclude that

$$\frac{\partial f}{\partial x}(c_1, b) = \frac{\partial f}{\partial x}(a, b) + \epsilon_1 \quad \text{and} \quad \frac{\partial f}{\partial y}(a+\Delta x, c_2) = \frac{\partial f}{\partial y}(a, b) + \epsilon_2, \quad (16.3.5)$$

where both ϵ_1 and ϵ_2 approach 0 as Δx and Δy approach 0.

Combining (16.3.4) and (16.3.5) gives the key to estimating the change in the function f. We state this important result as a theorem.

Theorem 16.3.1. *Let f have continuous partial derivatives f_x and f_y for all points within some disk with center at the point (a, b). Then Δf, which is the change $f(a+\Delta x, b+\Delta y) - f(a, b)$, can be written*

$$\Delta f = \frac{\partial f}{\partial x}(a, b)\Delta x + \frac{\partial f}{\partial y}(a, b)\Delta y + \epsilon_1\Delta x + \epsilon_2\Delta y, \qquad (16.3.6)$$

where ϵ_1 and ϵ_2 approach 0 as Δx and Δy approach 0. (Both ϵ_1 and ϵ_2 are functions of the four variables a, b, Δx and Δy.)

This equation is the core of this section.

The term $f_x(a, b)\Delta x$ estimates the change due to the change in the x-coordinate, while $f_y(a, b)\Delta y$ estimates the change due to the change in the y-coordinate.

We will call $f(x, y)$ **differentiable** at (a, b) if (16.3.6) holds. In particular if the partial derivatives f_x and f_y exist in a disk around (a, b) and are continuous at (a, b), then f is differentiable at (a, b).

Since ϵ_1 and ϵ_2 in (16.3.6) both approach 0 as Δx and Δy approach 0,

$$\Delta f \approx \frac{\partial f}{\partial x}(a, b)\Delta x + \frac{\partial f}{\partial y}(a, b)\Delta y, \qquad (16.3.7)$$

The approximation (16.3.7) gives us a way to estimate Δf when Δx and Δy are small.

EXAMPLE 1 Estimate $(2.1)^2(0.95)^3$.

SOLUTION Let $f(x,y) = x^2y^3$. We wish to estimate $f(2.1, 0.95)$. We know that $f(2,1)$ equals $2^2 1^3 = 4$. We use (16.3.7) to estimate $\Delta f = f(2.1, 0.95) - f(2,1)$. We have

$$\frac{\partial(x^2y^3)}{\partial x} = 2xy^3 \quad \text{and} \quad \frac{\partial(x^2y^3)}{\partial y} = 3x^2y^2.$$

Then

$$\frac{\partial f}{\partial x}(2,1) = 4 \quad \text{and} \quad \frac{\partial f}{\partial y}(2,1) = 12.$$

Since $\Delta x = 0.1$ and $\Delta y = -0.05$, we have

$$\Delta f = 4(0.1) + 12(-0.05) = 0.4 - 0.6 = -0.2.$$

Thus $(2,1)^2(0.95)^3$ is approximately $4 + (-0.2) = 3.8$. ◇

The exact value is 3.78102375.

DOUG: You may want to do more with "approximation". SHERMAN: What do you mean by this?

The Chain Rule

We begin with two special cases of the chain rule for functions of more than one variable. Afterward we will state the chain rule for functions of any number of variables.

The first theorem considers the case when $z = f(xy)$ and x and y are functions of just one variable t. The second theorem is more general, where x and y may be functions of two variables, t and u.

Theorem. *Chain Rule – Special Case #1 Let $z = f(x,y)$ have continuous partial derivatives f_x and f_y, and let $x = x(t)$ and $y = y(t)$ be differentiable functions of t. Then z is a differentiable function of t and*

$$\frac{dz}{dt} = \frac{\partial z}{\partial x}\frac{dx}{dt} + \frac{\partial z}{\partial y}\frac{dy}{dt}. \qquad (16.3.8)$$

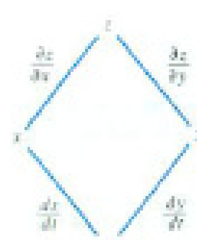
Proof

By definition,

$$\frac{dz}{dt} = \lim_{\Delta t \to 0} \frac{\Delta z}{\Delta t}.$$

Now, Δt induces changes Δx and Δy in x and y, respectively. According to Theorem 16.3.1,

$$\Delta z = \frac{\partial f}{\partial x}(x,y)\Delta x + \frac{\partial f}{\partial y}(x,y)\Delta y + \epsilon_1 \Delta x + \epsilon_2 \Delta y,$$



where $\epsilon_1 \to 0$ and $\epsilon_2 \to 0$ as Δx and Δy approach 0. (Keep in mind that x and y are fixed.) Thus

$$\frac{\Delta z}{\Delta t} = \frac{\partial f}{\partial x}(x,y)\frac{\Delta x}{\Delta t} + \frac{\partial f}{\partial y}(x,y)\frac{\Delta y}{\Delta t} + \epsilon_1\frac{\Delta x}{\Delta t} + \epsilon_2\frac{\Delta y}{\Delta t}.$$

and

$$\frac{dz}{dt} = \lim_{\Delta t \to 0}\frac{\Delta z}{\Delta t} = \frac{\partial f}{\partial x}(x,y)\frac{dx}{dt} + \frac{\partial f}{\partial y}(x,y)\frac{dy}{dt} + 0\frac{dx}{dt} + 0\frac{dy}{dt}.$$

This proves the theorem. •

> MEMORY AID: Each path produces one summand. And, each leg in each path produces one factor in that summand.

The two summands on the right-hand sides of (16.3.8) remind us of the chain rule for functions of one variable. Why is there a "+" in (16.3.8)? The "+" first appears in (16.3.4) and you can trace it back to Figure 16.3.1.

The diagram in Figure 16.3.2 helps in using this special case of the chain rule . There are two paths from the top variable z down to the bottom variable t. Label each edge with the appropriate partial derivative (or derivative). For each path there is a summand in the chain rule. The left-hand path (see Figure 16.3.3) gives us the summand

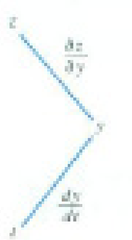


Figure 16.3.3:

$$\frac{\partial z}{\partial x}\frac{dx}{dt}.$$

The right-hand path (see Figure 16.3.4) gives us the summand

$$\frac{\partial z}{\partial y}\frac{dy}{dt}.$$

Then dz/dt is the sum of those two summands.

EXAMPLE 2 Let $z = x^2y^3$, $x = 3t^2$, and $y = t/3$. Find dz/dt when $t = 1$.
SOLUTION In order to apply the special case of the chain rule, compute z_x, z_y, dx/dt, and dy/dt:

$$\frac{\partial z}{\partial x} = 2xy^3 \qquad \frac{\partial z}{\partial y} = 3x^2y^2$$

$$\frac{dx}{dt} = 6t \qquad \frac{dy}{dt} = \frac{1}{3}.$$

By the special case of the chain rule,

$$\frac{dz}{dt} = 2xy^3 \cdot 6t + 3x^2y^2 \cdot \frac{1}{3}.$$

In particular, when $t = 1$, x is 3 and y is $\frac{1}{3}$. Therefore, when $t = 1$,

$$\frac{dz}{dt} \quad = \quad 2 \cdot 3 \left(\frac{1}{3}\right)^3 6 \cdot 1 + 3 \cdot 3^2 \left(\frac{1}{3}\right)^2 \frac{1}{3} = \frac{36}{27} + \frac{27}{27} = \frac{7}{3}.$$

◇

In Example 2, the derivative dz/dt can be found without using the theorem. To do this, express z explicitly in terms of t:

$$z = x^2y^3 = (3t^2)^2 \left(\frac{t}{3}\right)^3 = \frac{t^7}{3}.$$

Then

$$\frac{dz}{dt} = \frac{7t^6}{3}.$$

When $t = 1$, this gives

$$\frac{dz}{dt} = \frac{7}{3},$$

in agreement with the first computation.

EXAMPLE 3 The temperature at the points (x, y) on a window is $T(x, y)$. A bug wandering on the window is at the point $(x(t), y(t))$ at time t. How fast does the bug observe that the temperature of the glass changes as he crawls about?

SOLUTION The bug is asking us to find dT/dt. The chain rule (16.3.8) tells us that

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

The bug can influence this rate by crawling faster or slower. He may want to know the direction he should choose in order to cool off as quickly as possible. But we will not be able to tell him how to do this until the next section, Section 16.4. ◇

The proof of the next chain rule is almost identical to the proof of Theorem 16.3. (See Exercise 24.)

Theorem. *Chain Rule – Special Case #2 Let $z = f(x, y)$ have continuous partial derivatives, f_x and f_t. Let $x = x(t, u)$ and $y = (t, u)$ have continuous partial derivatives*

$$\frac{\partial x}{\partial t}, \qquad \frac{\partial x}{\partial u}, \qquad \frac{\partial y}{\partial t}, \qquad \frac{\partial y}{\partial u}.$$

Then

$$\frac{\partial z}{\partial t} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial t} \qquad \textit{and} \qquad \frac{\partial z}{\partial u} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial u}.$$

The variables are listed in Figure 16.3.5.

To find z_t, draw all the paths from z down to t. Label the edges by the appropriate partial derivative, as shown in Figure 16.3.6.

Each path from the top variable down to the bottom variable contributes a summand in the chain rule. The only difference between Figure 16.3.2 and Figure 16.3.6 is that ordinary derivatives dx/dt and dy/dt appear in Figure 16.3.2, while partial derivatives x_t and y_t appear in Figure 16.3.6.

In the first special case of the chain rule there are two middle variables and one bottom variable. In the second chain rule there are two middle variables and two bottom variables. The chain rule holds for any number of middle variables and any number of bottom variable. For instance, there may be three middle variables and, say, four bottom variables. In that case there are three summands for each of four partial derivatives.

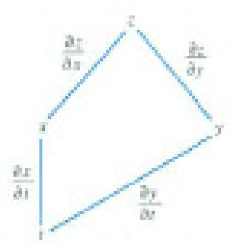


Figure 16.3.6:

In the next example there is only one middle variable and two bottom variables.

EXAMPLE 4 Let $z = f(u)$ be a function of a single variable. Let $u = 2x + 3y$. Then z is a composite function of x and y. Show that

$$2\frac{\partial z}{\partial y} = 3\frac{\partial z}{\partial x}. \tag{16.3.9}$$

SOLUTION We will evaluate both z_x and z_y by the chain rule and then check whether (16.3.9) is true.

To find z_x we consider all paths from z down to x. There is only one middle variable so there is only one path. Since $u = 2x + 3y$, $u_x = 2$. Thus

$$\frac{\partial z}{\partial x} = \frac{dz}{du}\frac{\partial u}{\partial x} = \frac{dz}{du}\cdot 2 = 2\frac{dz}{du}$$

(Note that one derivative is ordinary, while the other is a partial derivative.)

Next we find z_y. Again, there is only one summand. Since $u = 2x + 3y$, $u_y = 3$. Thus

$$\frac{\partial z}{\partial y} = \frac{dz}{du}\frac{\partial u}{\partial y} = \frac{dz}{du}\cdot 3 = 3\frac{dz}{du}.$$

Thus $z_x = 2dz/du$ and $z_y = 3dz/du$. Substitute these into the equation

$$2\frac{\partial z}{\partial y} = 3\frac{\partial z}{\partial x}$$

to see whether we obtain a true equation:

$$2\left(3\frac{dz}{du}\right) = 3\left(2\frac{dz}{du}\right). \tag{16.3.10}$$

Since (16.3.10) is true, we have verified (16.3.9). ◇

An Important Use of the Chain Rule

There is a fundamental difference between Example 2 and Example 4. In the first example, we were dealing with explicitly given functions. We did not really need to use the chain rule to find the derivative, dz/dt. As remarked after the example, we could have shown that $z = t^7/3$ and easily found that $dz/dt = 7t^6/3$. But in Example 4, we were dealing with a general type of function formed in a certain way: We showed that (16.3.9) holds for *every* differentiable function $f(u)$. No matter what $f(u)$ we choose, we know that $2z_y = 3z_x$.

Example 4 shows why the chain rule is important. It enables us to make *general statements* about the partial derivatives of an infinite number of functions, all of which are formed the same way. The next example illustrates this use again.

D'Alembert in 1746 obtained the partial differential equation for a vibrating string:

The wave equation also appears in the study of sound or light.

$$\frac{\partial^2 y}{\partial t^2} = k^2\frac{\partial^2 y}{\partial x^2}. \tag{16.3.11}$$

(See Figure C.21.3 in the CIE about the Wave in a Rope.) This "wave equation" created a great deal of excitement, especially since d'Alembert showed that any differentiable function of the form

$$g(x+kt) + h(x-kt)$$

is a solution.

Before we show that d'Alembert is right, we note that it is enough to check it for $g(x+kt)$. If you replace k by $-k$ in it, you will also have a solution since replacing k by $-k$ in (16.3.11) doesn't change the equation.

EXAMPLE 5 Show that any function $y = g(x+kt)$ satisfies the partial differential equation (16.3.11).

SOLUTION In order to find the partial derivatives y_{xx} and y_{tt} we express $y = g(x+kt)$ as a composition of functions:

$$y = g(u) \qquad \text{where} \qquad u = x + kt.$$

Note that g is a function of just one variable. Figure 16.3.7 lists the variables.

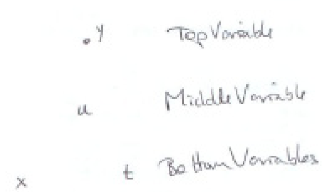


Figure 16.3.7:

Recall that $u = x + kt$.

We will compute y_{xx} and y_{tt} in terms of derivatives of g and then check whether (16.3.11) holds. We first compute y_{xx}. First of all,

$$\frac{\partial y}{\partial x} = \frac{dy}{du}\frac{\partial u}{\partial x} = \frac{dy}{du} \cdot 1 = \frac{dy}{du}. \tag{16.3.12}$$

(There is only one path from y down to x. See Figure 16.3.7.) In (16.3.12) dy/du is viewed as a function of x and t; that is, u is replaced by $x + kt$. Next,

$$\frac{\partial^2 y}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\partial y}{\partial x}\right) = \frac{\partial}{\partial x}\left(\frac{dy}{du}\right).$$

Now, dz/du, viewed as a function of x and t, may be expressed as a composite function. Letting $w = dy/du$, we have

$$w = f(u), \qquad \text{where} \qquad u = x + kt.$$

Therefore

$$\begin{aligned} \frac{\partial^2 y}{\partial x^2} &= \frac{\partial}{\partial x}\left(\frac{\partial y}{\partial x}\right) = \frac{\partial w}{\partial x} \\ &= \frac{dw}{du} \cdot \frac{\partial u}{\partial x} && \text{(only one path down to } x\text{)} \\ &= \frac{d}{du}\left(\frac{dy}{du}\right)\frac{\partial u}{\partial x} = \frac{d^2 y}{du^2} \cdot 1; \end{aligned}$$

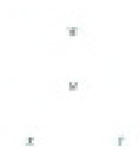


Figure 16.3.8:

hence

$$\frac{\partial^2 y}{\partial x^2} = \frac{d^2 y}{du^2}. \tag{16.3.13}$$

Then we also express y_{tt} in terms of d^2y/du^2, as follows. First of all,

$$\frac{\partial y}{\partial t} = \frac{dy}{du}\frac{\partial u}{\partial t} = \frac{dy}{du} \cdot k = k\frac{dy}{du}.$$

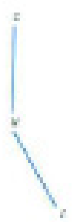


Figure 16.3.9:

(See Figure 16.3.9.)

Then

$$\begin{aligned} \frac{\partial^2 y}{\partial t^2} &= \frac{\partial}{\partial t}\left(\frac{\partial y}{\partial t}\right) = \frac{\partial}{\partial t}\left(k\frac{dy}{du}\right) \\ &= k\frac{d}{du}\left(\frac{dy}{du}\right) \cdot \frac{\partial u}{\partial t} && \text{(only one path down to } t\text{)} \\ &= k\frac{d^2 y}{du^2} \cdot k; \end{aligned}$$

hence

$$\frac{\partial^2 y}{\partial t^2} = k^2 \frac{d^2 z}{du^2} \tag{16.3.14}$$

Comparing (16.3.13) and (16.3.14) shows that

$$\frac{\partial^2 z}{\partial t^2} = k^2 \frac{d^2 z}{dx^2}$$

◇

Summary

The section opened by showing that under suitable assumptions on $f(x, y)$

$$\Delta f = \frac{\partial f}{\partial x}(a, b)\Delta x + \frac{\partial f}{\partial y}(a, b) + \epsilon_1 \Delta x + \epsilon_2 \Delta y, \qquad (16.3.15)$$

where ϵ_1 and ϵ_2 approach 0 as Δx and Δy approach 0. This gave us a way to estimate Δf, namely

$$\Delta f \approx \frac{\partial f}{\partial x}(a, b)\Delta x + \frac{\partial f}{\partial y}(a, b)\Delta y$$

"The change is due to both the change in x and the change in y." (16.3.15) generalizes to any number of variables and also is the basis for the various chain rules for partial derivatives. This is the general case:

If z is a function of $x_1, x_2, \ldots x_m$ and each x_i is a function of $t_1, t_2 \ldots t_n$, then there are n partial derivatives of $\partial z/\partial t_j$. Each is a sum of m products of the form $(\partial z/\partial x_i)(\partial x_i/\partial t_j)$. To do the bookkeeping, first make a roster as shown in Figure 16.3.10. To compute $\partial z/\partial t_j$, list all paths from z down to t_j, as shown in Figure 16.3.11. Each path that starts at z and goes down to t_j "contributes" a product. You do not have to be a great mathematician to apply the chain rule. However, you must do careful bookkeeping. First, display the top, middle, and bottom variables. Second, keep in mind that the number of middle variables determines the number of summands.

Some advice

Figure 16.3.10:

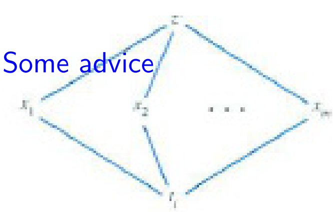


Figure 16.3.11:

EXERCISES for Section 16.3 *Key:* R–routine, M–moderate, C–challenging

In Exercises 1 to 4 verify the chain rule (Special Case #1, on page 1097) by computing dz/dt two ways: (a) with the chain rule, (b) without the chain rule (by writing z as a function of t).

1.[R] $z = x^2y^3$, $x = t^2$, $y = t^3$

2.[R] $z = xe^y$, $x = t$, $y = 1 + 3t$

3.[R] $z = \cos(xy^2)$, $x = e^{2t}$, $y = \sec(3t)$

4.[R] $z = \ln(x + 3y)$, $x = t^2$, $y = \tan(3t)$.

In Exercises 5 and 6 verify the chain rule (Special Case #2, on page 1099) by computing dz/dt two ways: (a) with the chain rule, (b) without the chain rule (by writing z as a function of t and u).

5.[R] $z = x^2y$, $x - 3t + 4u$, $y = 5t - u$

6.[R] $z = \sin(x + 3y)$, $x = \sqrt{t/u}$, $y = \sqrt{t} + \sqrt{u}$

7.[R] Assume that $z = f(x_1, x_2, x_3, x_4, x_5)$ and that each x_i is a function of t_1, t_2, t_3.

(a) List all variables, showing top, middle, and bottom variables.

(b) Draw the paths involved in expressing $\partial z/\partial t_3$ in terms of the chain rule.

(c) Express $\partial z/\partial t_3$ in terms of the sum of products of partial derivatives.

(d) When computing $\partial z/\partial t_2$, which variables are constant?

(e) When computing $\partial z/\partial t_3$, which variables are constant?

8.[R] If $z = f(g(t_1, t_2, t_3), h(t_1, t_2, t_3))$

(a) How many middle variables are there?

(b) How many bottom variables?

(c) What does the chain rule say about $\partial z/\partial t_3$? (Include a diagram showing the paths.)

9.[R] Find dz/dt if $z_x = 4$, $x_y = 3$, $dx/dt = 4$, and $dy/dt = 1$.

10.[R] Find dz/dt if $z_x = 3$, $z_y = 2$, $dx/dt = 4$, and $dy/dt = -3$.

11.[R] Let $z = f(x, y)$, $x = u + v$, and $y = u - v$.

(a) Show that $(z_x)^2 - (z_y)^2 = (z_u)(z_v)$. (Include diagrams.)

(b) Verify (a) when $f(x, y) = x^2 + 2y^3$.

12.[R] Let $z = f(x, y)$, $x = u^2 - v^2$, and $y = v^2 - u^2$.

(a) Show that
$$u\frac{\partial z}{\partial v} + v\frac{\partial z}{\partial u} = 0.$$
(Include diagrams.)

(b) Verify (a) when $f(xy) = \sin(x + 2y)$.

13.[R] Let $z = f(t - u, -t + u)$.

(a) Show that $\frac{\partial z}{\partial t} + \frac{\partial z}{\partial u} = 0$ (Include diagrams.)

(b) Verify (a) when $f(x, y) = x^2y$

14.[R] Let $w = f(x - y, y - z, z - x)$.

(a) Show that $\frac{\partial w}{\partial x} + \frac{\partial w}{\partial y} + \frac{\partial w}{\partial z} = 0$. (Include diagrams.)

(b) Verify (a) in the case $f(s, t, u) = s^2 + t^2 - u$.

15.[R] Let $z = f(u, v)$ where $u = ax+by$, $v = cx+dy$, and a, b, c, d are constants. Show that

(a) $\dfrac{\partial^2 z}{\partial x^2} = a^2\dfrac{\partial^2 f}{\partial u^2} + 2ac\dfrac{\partial^2 f}{\partial u \partial v} + c^2\dfrac{\partial^2 f}{\partial v^2}$

(b) $\dfrac{\partial^2 z}{\partial y^2} = b^2\dfrac{\partial^2 f}{\partial u^2} + 2bd\dfrac{\partial^2 f}{\partial u \partial v} + d^2\dfrac{\partial^2 f}{\partial v^2}$

(c) $\dfrac{\partial^2 z}{\partial x \partial y} = ab\dfrac{\partial^2 f}{\partial u^2} + (ad + bc)\dfrac{\partial^2 f}{\partial u \partial v} + cd\dfrac{\partial^2 f}{\partial v^2}.$

16.[R] Let a, b, and c be given constants and consider the partial differential equation

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

Assume a solution of the form $z = f(y + mx)$, where m is a constant. Show that for this function to be a solution, $am^2 + bm + c$ must be 0.

17.[R]

(a) Show that any function of the form $z = f(x+y)$ is a solution of the partial differential equation

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

(b) Verify (a) for $z = (x + y)^3$.

18.[R] Let $u(x, t)$ be the temperature at point x along a rod at time t. The function u satisfies the one-dimensional heat equation for a constant k:

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

(a) Show that $u(x, t) = e^{kt}g(x)$ satisfies the heat equation if $g(x)$ is any function such that $g''(x) = g(x)$.

(b) Show that if $g(x) = 3e^{-x} + 4e^{x}$, then $g''(x) = g(x)$.

19.[R]

(a) Show that any function of the form $z = f(x + y) + e^y f(x - y)$ is a solution of the partial differential equation

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

(b) Check (a) for $z = (x + y)^2 + e^y \sin(x - y)$.

20.[R] Let $z = f(x, y)$ denote the temperature at the point (x, y) in the first quadrant. If polar coordinates are used, then we would write $z = f(r, \theta)$.

(a) Express z_r in terms of z_x and x_y. HINT: What is the relation between rectangular coordinates (x, y) and polar coordinates (r, θ)?

(b) Express z_θ in terms of z_x and z_y.

(c) Show that

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

21.[R] Let $u = f(r)$ and $r = (x^2 + y^2 + z^2)^{1/2}$. Show that

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = \frac{d^2 u}{dr^2} + \frac{2}{r}\frac{du}{dr}.$$

22.[R] At what rate is the volume of a rectangular box changing when its width is 3 feet and increasing at the rate of 2 feet per second, its length is 8 feet and decreasing at the rate of 5 feet per second, and its height is 4 feet and increasing at the the rate of 2 feet per second?

23.[R] The temperature T at (x, y, z) in space is

$f(x, y, z)$. An astronaut is traveling in such a way that his x and y coordinates increase at the rate of 4 miles per second and his z coordinate decreases at the rate of 3 miles per second. Compute the rate dT/dt at which the temperature changes at a point where

$$\frac{\partial T}{\partial x} = 4, \quad \frac{\partial T}{\partial y} = 7, \quad \text{and} \quad \frac{\partial T}{\partial z} = 9.$$

24.[M] We proved Special Case #1 of the chain rule (page 1097), when there are two are two middle variables and one bottom variable. Prove Special Case #2 of the chain rule (page 1099), where there are two middle variables and two bottom variables.

25.[M] To prove the general chain rule when there are three middle variables, we need an analog of Theorem 16.3.1 concerning Δf when f is a function of three variables.

(a) Let $y = f(x, y, x)$ be a function of three variables. Show that

$$\begin{aligned} \Delta f &= f(x+\Delta x, y+\Delta y, z+\Delta z) - f(x,y,z) \\ &= (f(x+\Delta x, y, z) - f(x,y,z)) + (f(x+\Delta x, y+\Delta y, z) - f(x+\Delta x, y, z)) \\ &\quad +(f(x+\Delta x, y+\Delta y, z+\Delta z) - f(x+\Delta x, y+\Delta y, z)). \end{aligned}$$

(b) Using (a) show that

$$\Delta f = \frac{\partial f}{\partial x}(x,y,z)\Delta x + \frac{\partial f}{\partial y}(x,y,z)\Delta y + \frac{\partial f}{\partial z}(x,y,z)\Delta z + \epsilon_1 \Delta x + \epsilon_2 \Delta y + \epsilon_3 \Delta z,$$

where ϵ_1, ϵ_2, $\epsilon_3 \to 0$ as Δx, Δy, $\Delta z \to 0$.

(c) Obtain the general chain rule in the case of three middle variables and any number of bottom variables.

26.[M] Let $z = f(x, y)$, where $x = r\cos(\theta)$ and $y = r\sin(\theta)$. Show that

$$\frac{\partial^2 z}{\partial r^2} = \cos^2(\theta)\frac{\partial^2 f}{\partial x^2} + 2\cos(\theta)\sin(\theta)\frac{\partial^2 f}{\partial x \partial y} + \sin^2(\theta)\frac{\partial^2 f}{\partial y^2}.$$

27.[M] Let $u = f(x, y)$, where $x = r\cos(\theta)$ and $y = r\sin(\theta)$. Verify the following equation, which appears in electromagnetic theory,

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial u}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2 u}{\partial \theta^2} = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}.$$

28.[M] Let u be a function of x and y, where x and y are both functions of s and t. Show that

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

29.[C] Let (r, θ) be polar coordinates for the point (x, y) given in rectangular coordinates.

(a) From the relation $r = \sqrt{x^2 + y^2}$, show that $\partial r/\partial x = \cos(\theta)$.

(b) From the relation $r = x/\cos\theta$, show that $\partial r/\partial x = 1/\cos(\theta)$.

(c) Explain why (a) and (b) are not contradictory.

30.[C] In developing (16.3.6), we used the path that started at (x, y), went to $(x + \Delta x, y)$, and ended at $(x + \Delta x, y + \Delta y)$. Could we have used the path from (x, y), through $(x, y + \Delta y)$, to $(x + \Delta x, y + \Delta y)$ instead? If "no", explain why. If "yes," write out the argument, using the path.

In Exercises 31 to 34 concern homogeneous functions. A function $f(x, y)$ is homogeneous of degree r if $f(kx, ky) = k^r f(x, y)$ for all $k > 0$.

31.[R] Verify that each of the following functions is homogeneous of degree 1 and also verify that each satisfies the conclusion of Euler's theorem (with $r = 1$):

$$f(x, y) = x\frac{\partial f}{\partial x} + y\frac{\partial f}{\partial y}.$$

(a) $f(x, y) = 3x + 4y$

(b) $f(x, y) = x^3y^{-2}$

(c) $f(x, y) = xe^{x/y}$

32.[M] Show that each of the following functions is homogeneous, and find the degree r.

(a) $f(x, y) = x^2(\ln x - \ln y)$

(b) $f(x, y) = 1/\sqrt{x^2 + y^2}$

(c) $f(x, y) = \sin\left(\frac{y}{x}\right)$

33.[C] (See Exercise 31.) Show that if f is homogeneous of degree r, then $xf_x + yf_y = rf$. This is the general form of Euler's theorem.

34.[C] (See Exercise 33.) Verify Euler's theorem for each of the functions in Exercise 32.

35.[C] (See Exercise 32.) Show that if f is homogeneous of degree r, then $\partial f/\partial x$ is homogeneous of degree $r - 1$.

16.4 Directional Derivatives and the Gradient

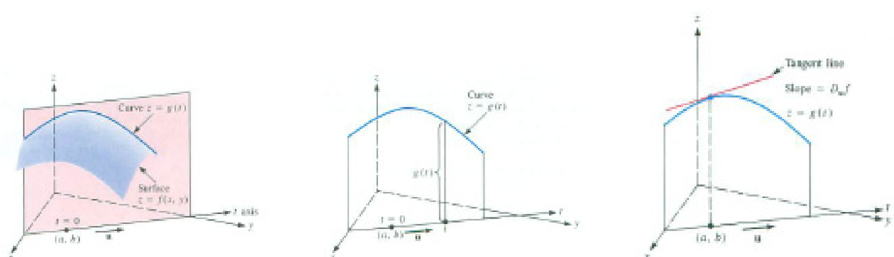
In this section we generalize the notion of a partial derivative to that of a directional derivative. Then we introduce a vector, called "the gradient," to provide a short formula for the directional derivative. The gradient will have other uses later in this chapter and in Chapter 18.

Directional Derivatives

If $z = f(x, y)$, the partial derivative $\partial f/\partial x$ tells us how rapidly z changes as we move the input point (x, y) in a direction parallel to the x-axis. Similarly, f_y tells how fast z changes as we move parallel to the y-axis. But we can ask, "How rapidly does z change when we move the input point (x, y) in any fixed direction in the xy plane?" The answer is given by the directional derivative.

It is important to remember that $\|\mathbf{u}\| = 1$.

Consider a function $z = f(x, y)$, let's say the temperature at (x, y). Let (a, b) be a point and let $\mathbf{u}$ be a unit vector in the xy plane. Draw a line through (a, b) and parallel to $\mathbf{u}$. Call it the t-axis and let its positive part point in the direction of $\mathbf{u}$. Place the 0 of the t-axis at (a, b). (See Figure 16.4.1.) Each value of t determines a point (x, y) on the t-axis and thus a value of z. Along the t-axis, z can therefore be viewed as a function of t, $z = g(t)$. The derivative dg/dt, evaluated at $t = 0$, is called the **directional derivative** of $z = f(x, y)$ at (a, b) in the direction $\mathbf{u}$. It is denoted $D_{\mathbf{u}}f$. The directional derivative is the slope of the tangent line to the curve $z = g(t)$ at $t = 0$. (See Figure 16.4.1(c).)



(a) (b) (c)

Figure 16.4.1: ARTIST: Improved figures are needed here.

When $\mathbf{u} = \mathbf{i}$, we obtain the directional derivative $D_{\mathbf{u}}f$, which is simply fx. When $\mathbf{u} = \mathbf{j}$, we obtain $D_{\mathbf{j}}f$, which is f_y.

The directional derivative generalizes the two partial derivatives f_x and f_y. After all, we can ask for the rate of change of $z = f(x, y)$ in any direction in the xy plane, not just the directions indicated by the vectors $\mathbf{i}$ and $\mathbf{j}$.

The following theorem shows how to compute a directional derivative.

Theorem. *(Directional Derivatives) If $f(x, y)$ has continuous partial derivatives f_x and f_y, then the directional derivative of f at (a, b) in the direction of* $\mathbf{u} = \cos(\theta)\mathbf{i} + \sin(\theta)\mathbf{j}$ *where θ is the angle between* $\mathbf{u}$ *and* $\mathbf{i}$ *is*

$$\frac{\partial f}{\partial x}(a, b)\cos(\theta) + \frac{\partial f}{\partial y}(a, b)\sin(\theta). \qquad (16.4.1)$$

Proof

The directional derivative of f at (a, b) in the direction $\mathbf{u}$ is the derivative of the function

$$g(t) = f(a + t\cos(\theta), b + t\sin(\theta))$$

when $t = 0$. (See Figure 16.3.2 and Figure 16.3.3.)

Now, g is a composite function

$$g(t) = f(x, y) \qquad \text{where} \qquad \begin{cases} x = a + t\cos(\theta) \\ y = b + t\sin(\theta). \end{cases}$$

The chain rule tells us that

$$g'(t) = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt}$$

Moreover,

$$\frac{dx}{dt} = \cos(\theta) \qquad \text{and} \qquad \frac{dy}{dt} = \sin(\theta).$$

Thus

$$g'(0) = \frac{\partial f}{\partial x}(a, b)\cos\theta + \frac{\partial f}{\partial y}(a, b)\sin\theta,$$

and the theorem is proved. •

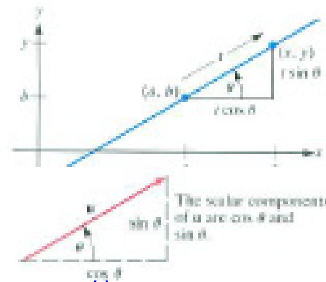


Figure 16.4.3:

Check (16.4.1) when $\theta = 0$

When $\theta = 0$, that is, $\mathbf{u} = \mathbf{i}$, (16.4.1) becomes

$$\frac{\partial f}{\partial x}(a, b)\cos(0) + \frac{\partial f}{\partial y}(a, b)\sin(0) = \frac{\partial f}{\partial x}(a, b)(1) + \frac{\partial f}{\partial y}(a, b)(0) = \frac{\partial f}{\partial x}(a, b).$$

Check (16.4.1) when $\theta = \pi$

When $\theta = \pi$, that is, $\mathbf{u} = -\mathbf{i}$, (16.4.1) becomes

$$\frac{\partial f}{\partial x}(a, b)\cos(\pi) + \frac{\partial f}{\partial y}(a, b)\sin(\pi) = \frac{\partial f}{\partial x}(a, b)(-1) + \frac{\partial f}{\partial y}(a, b)(0) = -\frac{\partial f}{\partial x}(a, b).$$

(This makes sense: If the temperature increases as you walk east, then it decreases when you walk west.)

Check (16.4.1) when $\theta = \frac{\pi}{2}$

When $\theta = \frac{\pi}{2}$, that is, $\mathbf{u} = \mathbf{j}$, (16.4.1) asserts that the directional derivative is

$$\frac{\partial f}{\partial x}(a, b)\cos(\frac{\pi}{2}) + \frac{\partial f}{\partial y}(a, b)\sin(\frac{\pi}{2}) = \frac{\partial f}{\partial x}(a, b)(0) + \frac{\partial f}{\partial y}(a, b)(1) = \frac{\partial f}{\partial y}(a, b).$$

which also is expected.

EXAMPLE 1 Compute the derivative of $f(x, y) = x^2y^3$ at $(1, 2)$ in the direction given by the angle $\pi/3$. (That is, $\mathbf{u} = \cos(\pi/3)\mathbf{i} + \sin(\pi/3)\mathbf{j}$.) Interpret the results if f describes a temperature distribution.

SOLUTION First of all,

$$\frac{\partial f}{\partial x} = 2xy^3 \quad \text{and} \quad \frac{\partial f}{\partial y} = 3x^2y^2.$$

Hence

$$\frac{\partial f}{\partial x}(1, 2) = 16 \quad \text{and} \quad \frac{\partial f}{\partial y}(1, 2) = 12.$$

Second,

$$\cos\left(\frac{\pi}{3}\right) = \frac{1}{2} \quad \text{and} \quad \sin\left(\frac{\pi}{3}\right) = \frac{\sqrt{3}}{2}.$$

Thus the derivative of f in the direction given by $\theta = \pi/3$ is

$$16\left(\frac{1}{2}\right) + 12\left(\frac{\sqrt{3}}{2}\right) = 8 + 6\sqrt{3} \approx 18.3923.$$

If x^2y^3 is the temperature in degrees at the point (x, y), where x and y are measured in centimeters, then the rate at which the temperature changes at $(1, 2)$ in the direction given by $\theta = \pi/3$, is approximately 18.4 degrees per centimeter. ◇

The Gradient

Equation (16.4.1) resembles the formula for the dot product. To exploit this similarity, it is useful to introduce the vector whose scalar components are $f_x(a, b)$ and $f_y(a, b)$.

DEFINITION *(The gradient of $f(x, y)$.)* The vector

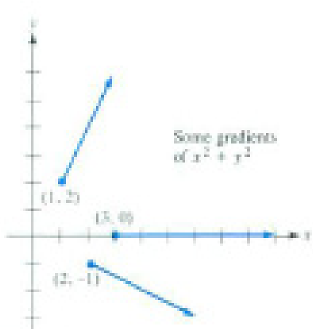
$$\frac{\partial f}{\partial x}(a, b)\mathbf{i} + \frac{\partial f}{\partial y}(a, b)\mathbf{j}$$

is the **gradient** of f at (a, b) and is denoted ∇f. (It is also called "del f," because of the upside-down delta ∇.)

The del symbol is in boldface to emphasize that the gradient of f is a vector. For instance, let $f(x, y) = x^2 + y^2$. We compute and draw ∇f at a few points, listed in the following table:

Figure 16.4.4 shows ∇f, with the tail of ∇f placed at the point where ∇f is computed.

In vector notation, Theorem 16.4 reads as follows:



(x, y)	$\frac{\partial f}{\partial x} = 2x$	$\frac{\partial f}{\partial y} = 2y$	∇f
$(1, 2)$	2	4	$2\mathbf{i} + 4\mathbf{j}$
$(3, 0)$	6	0	$6\mathbf{i}$
$(2, -1)$	4	-2	$4\mathbf{i} - 2\mathbf{j}$

Table 16.4.1:

Theorem. *Directional Derivative - Rephrased If $z = f(x, y)$ has continuous partial derivatives f_x and f_y, then at (a, b)*

$$D_{\mathbf{u}}f = \nabla f(a, b) \cdot \mathbf{u} = (f_x(a, b)\mathbf{i} + f_y(a, b)\mathbf{j}) \cdot \mathbf{u}.$$

The gradient is introduced not merely to simplify the computation of directional derivatives. Its importance is made clear in the next theorem.

A Different View of the Gradient

The gradient vector provides two important pieces of geometric information about a function. The gradient vector, $\nabla f(a, b)$, always points in the direction in which the function increases most rapidly from the point (a, b). In the same way, the negative of the gradient vector, $-\nabla f(a, b)$, always points in the direction in which the function decreases most rapidly from the point (a, b). And, the length of the gradient vector, $\|\nabla f(a, b)\|$, is the largest directional derivative of f at (a, b).

The meaning of $\|\nabla f\|$ and the direction of ∇f

Theorem. *Significance of ∇f Let $z = f(x, y)$ have continuous partial derivatives f_x and f_y. Let (a, b) be a point in the plane where ∇f is not $\mathbf{0}$. Then the length of ∇f at (a, b) is the largest directional derivative of f at (a, b). The direction of ∇f is the direction in which the directional derivative at (a, b) has its largest value.*

Proof

By the definition of the directional derivative, if $\mathbf{u}$ is a unit vector, then, at (a, b),

$$D_{\mathbf{u}}f = \nabla f \cdot \mathbf{u}.$$

By the definition of the dot product

$$\nabla f \cdot \mathbf{u} = \|\nabla f\|\|\mathbf{u}\| \cos(\alpha),$$

where α is the angle between ∇f and $\mathbf{u}$, as shown in Figure 16.4.5. Since $|\mathbf{u}| = 1$,

$$D_{\mathbf{u}}f = \|\nabla f\| \cos(\alpha). \qquad (16.4.2)$$

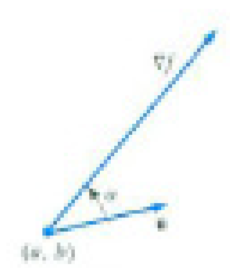


Figure 16.4.5:

The largest value of $\cos(\alpha)$ for $0 \le \alpha \le \pi$, occurs when $\cos(\alpha) = 1$; that is, when $\alpha = 0$. Thus, by (16.4.2), the largest directional derivative of $f(x, y)$ at (a, b) occurs when the direction is that of ∇f at (a, b). For that choice of $\mathbf{u}$, $D_{\mathbf{u}} f = \|\nabla f\|$. This proves the theorem. •

What does this theorem tell a bug wandering around on a flat piece of metal? If it is at the point (a, b) and wishes to get warmer as quickly as possible, it should compute the gradient of the temperature function and then go in the direction indicated by that gradient.

EXAMPLE 2 What is the largest direction derivative of $f(x, y) = x^2y^3$ at $(2, 3)$? In what direction does this maximum directional derivative occur?
SOLUTION At the point (x, y),

$$\nabla f = 2xy^3\mathbf{i} + 3x^2y^2\mathbf{j}.$$

Thus at $(2, 3)$,

$$\nabla f = 108\mathbf{i} + 108\mathbf{j},$$

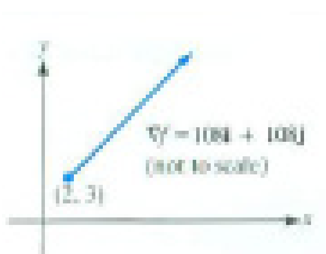


Figure 16.4.6:
Direction of fastest decrease is $-\nabla f$

which is sketched in Figure 16.4.6 (not to scale). Note that its angle θ is $\pi/4$. The maximal directional derivative of x^2y^3 at $(2, 3)$ is $\|\nabla f\| = 108\sqrt{2} \approx 152.735$. This is achieved at the angle $\theta = \pi/4$, relative to the x-axis, that is, for

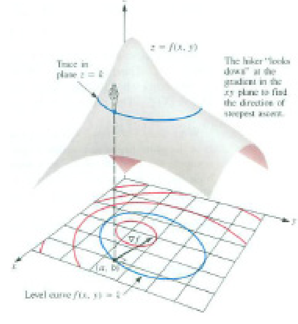
$$\mathbf{u} = \cos\left(\frac{\pi}{4}\right)\mathbf{i} + \sin\left(\frac{\pi}{4}\right)\mathbf{j} = \frac{\sqrt{2}}{2}\mathbf{i} + \frac{\sqrt{2}}{2}\mathbf{j}.$$

◇

Incidentally, if $f(x, y)$ denotes the temperature at (x, y), the gradient ∇f helps indicate the direction in which heat flows. It tends to flow "toward the coldest," which boils down to the mathematical assertion, "Heat tends to flow in the direction of $-\nabla f$."

The gradient and directional derivative have been interpreted in terms of a temperature distribution in the plane and a wandering bug. It is also instructive to interpret these concepts in terms of a hiker on the surface of a mountain.

Consider a mountain above the xy plane. The elevation of the point on the surface above the point $x, y)$ will be denoted by $f(x, y)$. The directional derivative $D_{\mathbf{u}} f$ indicates the rate at which altitude changes per unit change in *horizontal* distance in the direction of $\mathbf{u}$. The gradient ∇f at (a, b) points in the compass direction the hiker should choose to climb in the direction of the steepest ascent. The length of ∇f tells the hiker the steepest slope available. (See Figure 16.4.7.)



Generalization to $f(x, y, z)$

The notions of directional derivative and gradient can be generalized with little effort to functions of three (or more) variables. It is easiest to interpret the directional derivative of $f(x, y, z)$ in a particular direction in space as indicating the rate of change of the function in that direction in space. A useful interpretation is how fast the temperature changes in a given direction.

Let $\mathbf{u}$ be a unit vector in space, with direction angles α, β, and γ. Then $\mathbf{u} = \cos\alpha\mathbf{i} + \cos\beta\mathbf{j} + \cos\gamma\mathbf{k}$. We now define the derivative of $f(x, y, z)$ in the direction $\mathbf{u}$.

DEFINITION *(Directional Derivative of $f(x, y, z)$.)* The **directional derivative** of f at (a, b, c) in the direction of the unit vector $\mathbf{u} = \cos(\alpha)\mathbf{i} + \cos(\beta)\mathbf{j} + \cos(\gamma)\mathbf{k}$ is $g'(0)$, where g is defined by

$$g(t) = f(a + t\cos(\alpha), b + t\cos(\beta), c + t\cos(\gamma)).$$

It is denoted $D_{\mathbf{u}}f$.

Note that t is the measure of length along the line through (a, b, c) with direction angles α, β, and γ. Therefore $D_{\mathbf{u}}f$ is just a derivative along the t-axis.

The proof of the following theorem for a function of three variables is like those given earlier in this section for functions of two variables.

Theorem. *Directional Derivative of $f(x, y, z)$ If $f(x, y, z)$ has continuous partial derivatives f_x, f_y, and f_z, then the directional derivative of f at (a, b, c) in the direction of the unit vector $\mathbf{u} = \cos(\alpha)\mathbf{i} + \cos(\beta)\mathbf{j} + \cos(\gamma)\mathbf{k}$ is*

$$\frac{\partial f}{\partial x}(a, b, c)\cos(\alpha) + \frac{\partial f}{\partial y}(a, b, c)\cos(\beta) + \frac{\partial f}{\partial z}(a, b, c)\cos(\gamma).$$

DEFINITION *(The gradient of $f(x, y, z)$.)* The vector

$$\frac{\partial f}{\partial x}(a, b, c)\mathbf{i} + \frac{\partial f}{\partial y}(a, b, c)\mathbf{j} + \frac{\partial f}{\partial z}(a, b, c)\mathbf{k}$$

is the **gradient** of f at (a, b, c) and is denoted ∇f.

This theorem thus asserts that

> the derivative of $f(x, y, z)$ in the direction of the unit vector $\mathbf{u}$ equals the dot product of $\mathbf{u}$ and the gradient of f:
>
> $$D_{\mathbf{u}}f = \nabla f \cdot \mathbf{u}.$$

Just as in the case of a function of two variables, ∇f evaluated at (a, b, c), points in the direction $\mathbf{u}$ that produces the largest directional derivative at (a, b, c). Moreover $\|\nabla f\|$ is that largest directional derivative. Just as in the two variable case, the key steps in the proof of this theorem are writing $\nabla f \cdot \mathbf{u} = \|\nabla f\| \|\mathbf{u}\| \cos(\nabla f, \mathbf{u})$ and recalling that $\mathbf{u}$ is a unit vector.

EXAMPLE 3 The temperature at the point (x, y, z) in a solid piece of metal is given by the formula $f(x, y, z) = c^{2x+y+3z}$ degrees. In what direction at the point $(0, 0, 0)$ does the temperature increase most rapidly?

SOLUTION First compute

$$\frac{\partial f}{\partial x} = 2e^{2x+y+3z}, \quad \frac{\partial f}{\partial y} = e^{2x+y+3z}, \quad \frac{\partial f}{\partial z} = 3e^{2x+y+3z}.$$

Then form the gradient vector:

$$\nabla f = 2e^{2x+y+3z}\mathbf{i} + e^{2x+y+3z}\mathbf{j} + 3e^{2x+y+3z}\mathbf{k}.$$

At $(0, 0, 0)$,

$$\nabla f = 2\mathbf{i} + \mathbf{j} + 3\mathbf{k}.$$

Consequently, the direction of most rapid increase in temperature is that given by the vector $2\mathbf{i} + \mathbf{j} + 3\mathbf{k}$. The rate of increase is then

$$\|2\mathbf{i} + \mathbf{j} + 3\mathbf{k}\| = \sqrt{14} \text{ degrees per unit length.}$$

If the line through $(0, 0, 0)$ parallel to $2\mathbf{i} + \mathbf{j} + 3\mathbf{k}$ is given a coordinate system so that it becomes the t-axis, with $t = 0$ at the origin and the positive part in the direction of $2\mathbf{i} + \mathbf{j} + 3\mathbf{k}$, the $df/dt = \sqrt{14}$ at 0. ◇

The gradient was denoted Δ by Hamilton in 1846. By 1870 it was denoted ∇, an upside-down delta, and therefore called "atled." In 1871 Maxwell wrote, "The quantity ∇P is a vector. I venture, with much diffidence, to call it the *slope* of P." The name "slope" is no longer used, having been replaced by "gradient." "Gradient" goes back to the word "grade," the slope of a road or surface. The name "del" first appeared in print in 1901, in *Vector Analysis, A text-book for the use of students of mathematics and physics founded upon the lectures of J. Willard Gibbs*, by E.B. Wilson.

Summary

We defined the derivative of $f(x, y)$ at (a, b) in the direction of the unit vector $\mathbf{u}$ in the xy plane and the derivative of $f(x, y, z)$ at (a, b, c) in the direction of the unit vector $\mathbf{u}$ in space. Then we introduced the gradient vector ∇f in terms of its components and obtained the formula

$$D_{\mathbf{u}} f = \nabla f \cdot \mathbf{u}.$$

By examining this formula we saw that the length and direction of ∇f at a given point are significant:

- ∇f points in the direction $\mathbf{u}$ that maximizes $D_{\mathbf{u}} f$ at the given point
- $\|\nabla f\|$ is the maximum directional derivative of f at the given point.

EXERCISES for Section 16.4 *Key:* R–routine, M–moderate, C–challenging

As usual, we assume that all functions mentioned have continuous partial derivatives.

In Exercises 1 and 2 compute the directional derivatives of x^4y^5 at $(1,1)$ in the indicated directions.

1.[R] (a) $\mathbf{i}$, (b) $-\mathbf{i}$, (c) $\cos(\pi/3)\mathbf{i}+\sin(\pi/3)\mathbf{j}$
$\cos(\pi/4)\mathbf{i}+\sin(\pi/4)\mathbf{j}$

2.[R] (a) $\mathbf{j}$, (b) $-\mathbf{j}$, (c)

In Exercises 3 and 4 compute the directional derivatives of x^2y^3 in the directions of the given vectors.

3.[R] (a) $\mathbf{j}$, (b) $\mathbf{k}$, (c) $-\mathbf{i}$

4.[R] (a) $\mathbf{i}+\mathbf{j}+\mathbf{k}$, (b) $2\mathbf{i}-\mathbf{j}+2\mathbf{k}$, (c) $\mathbf{i}+\mathbf{k}$
NOTE: These are not unit vectors. First construct a unit vector with the same direction.

5.[R] Assume that, at the point $(2,3)$, $\partial f/\partial x = 4$ and $\partial f/\partial y = 5$.

(a) Draw ∇f at $(2,3)$.

(b) What is the maximal directional derivative of f at $(2,3)$?

(c) For which $\mathbf{u}$ is $D_{\mathbf{u}}f$ at $(2,3)$ maximal? (Write $\mathbf{u}$ in the form $x\mathbf{i}+y\mathbf{j}$.)

6.[R] Assume that, at the point $(1,1)$, $\partial f/\partial x = 3$ and $\partial f/\partial y = -3$.

(a) Draw ∇f at $(1,1)$.

(b) What is the maximal directional derivative of f at $(1,1)$?

(c) For which $\mathbf{u}$ is $D_{\mathbf{u}}f$ at $(1,1)$ maximal? (Write $\mathbf{u}$ in the form $x\mathbf{i}+y\mathbf{j}$.)

In Exercises 7 and 8 compute and draw ∇f at the indicated points for the given functions.

7.[R] $f(x,y) = x^2y$ at (a) $(2,5)$, (b) $(3,1)$ $1/\sqrt{x^2+y^2}$ at (a) $(1,2)$, (b) $(3,0)$

8.[R] $f(x,y) =$

9.[R] If the maximal directional derivative of f at (a,b) is 5, what is the minimal directional derivative there? Explain.

10.[R] For a given function $f(x,y)$ at a given point (a,b) is there always a direction in which the directional derivative is 0? Explain.

11.[R] If $(\partial f/\partial x)(a,b) = 2$ and $(\partial f/\partial t)(a,b) = 3$, in what direction should a directional derivative at (a,b) be computed in order that it be

(a) 0?

(b) as large as possible?

(c) as small as possible?

12.[R] If, at the point (a,b,c), $\partial f/\partial x = 2$, $\partial f/\partial y = 3$, $\partial f/\partial z = 4$, what is the largest directional derivative of f at (a,b,c)?

13.[R] Assume that $f(1,2) = 2$ and $f(0.99, 2.01) = 1.98$.

(a) Which directional derivatives $D_{\mathbf{u}}f$ at $(1,2)$ can be estimated with this information? (Give $\mathbf{u}$.)

(b) Estimate the directional derivatives in (a).

14.[R] Assume that $f(1,1,1) = 3$ and $f(1.1, 1.2, 1.1) = 3.1$.

(a) Which directional derivatives $D_{\mathbf{u}}f$ at $(1,1,1)$ can be estimated with this information? (Give $\mathbf{u}$.)

(b) Estimate the directional derivatives in (a).

15.[R] When a bug crawls east, it discovers that the temperature increases at the rate of 0.02° per centimeter. When it crawls north, the temperature decreases at the rate of −0.03° per centimeter.

(a) If the bug crawls south, at what rate does the temperature change?

(b) If the bug crawls 30° north of east, at what rate does the temperature change?

(c) If the bug is happy with its temperature, in what direction should it crawl to try to keep the temperature the same?

16.[R] A bird is very sensitive to the temperature. It notices that when it flies in the direction $\mathbf{i}$, the temperature increases at the rate of 0.03° per centimeter. When it flies in the direction $\mathbf{j}$, the temperature decreases at the rate of 0.02° per centimeter. When it flies in the direction $\mathbf{k}$ the temperature increases at the rate of 0.05° per centimeter. It decides to fly off in the direction of the vector $(2,5,1)$. Will it be getting warmer or colder?

17.[R] Assume that $f(1,2)=3$ and that the directional derivative of f at $(1,2)$ in the direction of the (nonunit) vector $\mathbf{i}+\mathbf{j}$ is 0.7. Use this information to estimate $f(1.1,2.1)$.

18.[R] Assume that $f(1,1,2)=4$ and that the directional derivative of f at $(1,1,2)$ in the direction of the vector from $(1,1,2)$ to $(1.01,1.02,1.99)$ is 3. Use this information to estimate $f(0.99,0.98,2.01)$.

In Exercises 19 to 24 find the directional derivative of the function in the given direction and the maximum directional derivative.

19.[R] xyz^2 at $(1,0,1)$; $\mathbf{i}+\mathbf{j}+\mathbf{k}$

20.[R] x^3yz at $(2,1,-1)$; $2\mathbf{i}-\mathbf{k}$

21.[R] $e^{xy\sin(z)}$ at $(1,1,\pi/4)$; $\mathbf{i}+\mathbf{j}++3\mathbf{k}$

22.[R] $\arctan(\sqrt{x^2+y+z})$ at $(1,1,1)$; $-\mathbf{i}$

23.[R] $\ln(1+xyz)$ at $(2,3,1)$; $-\mathbf{i}+\mathbf{j}$

24.[R] $x^xye^{z^2}$ at $(1,1,0)$; $\mathbf{i}-\mathbf{j}+\mathbf{k}$

25.[R] Let $f(x,y,z)=2x+3y+z$.

(a) Compute ∇f at $(0,0,0)$ and at $(1,1,1)$.

(b) Draw ∇f for the two points in (a), in each case putting its tail at the point.

26.[R] Let $f(x,y,z)=x^2+y^2+z^2$.

(a) Compute ∇f at $(2,0,0)$, $(0,2,0)$ and $(0,0,2)$.

(b) Draw ∇f for the three points in (a), in each case putting its tail at the point.

27.[M] Assume that ∇f at (a,b) is not $\mathbf{0}$. Show that there are two unit vectors $\mathbf{u}_1$ and $\mathbf{u}_2$, such that the directional derivatives of f at (a,b) in the direction of $\mathbf{u}_1$ and $\mathbf{u}_2$ are 0.

28.[M] Assume that ∇f at (a,b,c) is not $\mathbf{0}$. How many unit vectors $\mathbf{u}$ are there such that $D_{\mathbf{u}}f=0$? Explain.

29.[R] Let $T(x,y,z)$ be the temperature at the point (x,y,z). Assume that ∇T at $(1,1,1)$ is $2\mathbf{i}+3\mathbf{j}+4\mathbf{k}$.

(a) Find $D_{\mathbf{u}}T$ at $(1,1,1)$ if $\mathbf{u}$ is in the direction of the vector $\mathbf{i}-\mathbf{j}+2\mathbf{k}$.

(b) Estimate the change in temperature as you move from the point $(1,1,1)$ a distance 0.2 in the direction of the vector $\mathbf{i}-\mathbf{j}+2\mathbf{k}$.

(c) Find three unit vectors $\mathbf{u}$ such that $D_{\mathbf{u}}T=0$ at $(1,1,1)$.

30.[R] A bug at the point $(1, 2)$ is very sensitive to the temperature and observes that if it moves in the direction $\mathbf{i}$ the temperature increases at the rate of $2°$ per centimeter. If it moves in the direction $\mathbf{j}$, the temperature decreases at the rate of $2°$ per centimeter. In what direction should it move if it wants

(a) to warm up most rapidly?

(b) to cool off most rapidly?

(c) to change the temperature as little as possible?

31.[R] Let $f(x, y) = 1/\sqrt{x^2 + y^2}$; the function f is defined everywhere except at $(0, 0)$. Let $\mathbf{r} = \langle x, y\rangle$.

(a) Show that $\nabla f = -r/\|\mathbf{r}\|^3$.

(b) Show that $\|\nabla f\| = -1/\|\mathbf{r}\|^2$.

32.[R] Let $f(x, y, z) = 1/\sqrt{x^2 + y^2 + z^2}$, which is defined everywhere except at $(0, 0, 0)$. (This function is related to the potential in a gravitational field due to a point-mass.) Let $\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$. Express ∇f in terms of r.

s a typo
∇. What
want? I
$-1/c^2$.

33.[R] Let $f(x, y) = x^2 + y^2$. Prove that if (a, b) is an arbitrary point on the curve $x^2 + y^2 = 9$, then ∇f computed at (a, b) is perpendicular to the tangent line to that curve at (a, b).

34.[R] Let $f(x, y, z)$ equal temperature at (x, y, z). Let $P = (a, b, c)$ and Q be a point very near (a, b, c). Show that $\nabla f \cdot \overrightarrow{PQ}$ is a good estimate of the change in temperature from point P to point Q.

rcises 28
ilar, but
27 and)
ter, and
Or, one
Chapter
mmary?

35.[R]

(a) If $(\partial f/\partial x)(a, b, c) = 2$, $(\partial f/\partial y)(a, b, c) = 3$ and $(\partial f/\partial z)(a, b, c) = 1$, find three different unit vectors $\mathbf{u}$ such that $D_{\mathbf{u}}f$ at (a, b, c) is 0.

(b) How many unit vectors $\mathbf{u}$ are there such that $D_{\mathbf{u}}f$ at (a, b, c) is 0?

36.[C] Let $f(x, y) = xy$.

(a) Draw the level curve $xy = 4$ carefully.

(b) Compute ∇f at three convenient points on that level curve and draw it with its tail at the point where it is evaluated.

(c) What angle does ∇f seem to make with the curve at the point where it is evaluated?

(d) Prove that the angle is what you think it is.

37.[M] Let (x, y) be the temperature at (x, y). Assume that ∇f at $(1, 1)$ is $2\mathbf{i} + 3\mathbf{j}$. A bug is crawling northwest at the rate of 3 centimeters per second. Let $g(t)$ be the temperature at the point where the bug is at time t seconds. Then dg/dt is the rate at which temperature changes on the bug's journey (degrees per second.) Find dg/dt when the bug is at $(1, 1)$.

38.[R] If $f(P)$ is the electric potential at the point P, then the electric field $\mathbf{E}$ at P is given by $-1/c^2\nabla f$. Calculate $\mathbf{E}$ if $f(x, y) = \sin(\alpha x)\cos(\beta y)$, where α and β are constants.

39.[R] The equality $\partial^2 f/\partial x\partial y = \partial^2 f/\partial y\partial x$ can be written as $D_{\mathbf{i}}(D_{\mathbf{j}}f) = D_{\mathbf{j}}(D_{\mathbf{i}}f)$. Show for any two unit vectors $\mathbf{u}_1$ and $\mathbf{u}_2$ that $D_{\mathbf{u}_2}(D_{\mathbf{u}_1}f) = D_{\mathbf{u}_1}(D_{\mathbf{u}_2}f)$. (Assume that all partial derivatives of f of all orders are continuous.)

40.[C] Without the aid of vectors, prove that the maximum value of

$$g(\theta) = \partial f/\partial x(a, b)\cos(\theta) + \partial f/\partial y(a, b)\sin(\theta)$$

is $\sqrt{(\partial f/\partial x(a, b))^2 + (\partial f/\partial y(a, b))^2}$. NOTE: This is the first part of the theorem about the significance of the gradient, on page 1111.

41.[R] Figure 16.4.8 shows two level curves of a function $f(x, y)$ near the point $(1, 2)$, namely $f(x, y) = 3$ and $f(x, y) = 3.02$. Use the diagram to estimate

(a) $D_{\mathbf{i}}f$ at $(1, 2)$,

(b) $D_{\mathbf{j}}f$ at $(1, 2)$,

(c) Draw ∇f at $(1, 2)$.

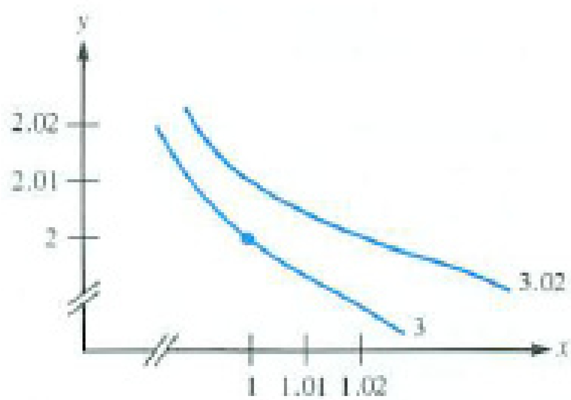


Figure 16.4.8:

42.[C] Why is a unit vector $\mathbf{u}$ in the xy-plane described by a single angle θ, but a unit vector in space is described by three angles?

43.[M] Let f and g be two vector functions defined throughout the xy-plane. Assume they have the same gradient, $\nabla f = \nabla g$. Must $f = g$? Is there any relation between f and g?

16.5 Normals and Tangent Planes

In this section we first find how to obtain a normal vector to a curve given implicitly, as a level curve $f(x, y) = k$. Then we find how to obtain a normal to a surface given implicitly, as a level surface $f(x, y, z) = k$. With the aid of this vector we define the tangent plane to a surface at a given point on the surface.

Normals to a Curve in the xy Plane

We saw in Section 14.4 how to find a normal vector to a curve when the curve is given parametrically, $\mathbf{r} = \mathbf{G}(t)$. Now we will see how to find a normal when the curve is given implicitly, as a level curve $f(x, y) = k$. Throughout this section we assume that the various functions are "well behaved." In particular, curves have continuous tangent vectors and functions have continuous partial derivatives.

Theorem. *The gradient ∇f at (a, b) is a normal to the level curve of f passing through (a, b).*

Proof
Let $\mathbf{G}(t) = x(t)\mathbf{i} + y(t)\mathbf{j}$ be a parameterization of the level curve of f that passes through the point (a, b). On this curve, $f(x, y)$ is a constant and has the value $f(a, b)$. Let $\mathbf{G}'(t_0)$ be the tangent vector to the curve at (a, b) and let the gradient of f at (a, b) be $\nabla f(a, b) = f_x(a, b,)\mathbf{i} - f_y(a, b)\mathbf{j}$. We wish to show that

$$\nabla f \cdot \mathbf{G}'(t_0) = 0;$$

that is,

$$\frac{\partial f}{\partial x}(a, b)\frac{dx}{dt}(t_0) + \frac{\partial f}{\partial y}(a, b)\frac{dy}{dt}(t_0) = 0. \qquad (16.5.1)$$

The left side of (16.5.1) has the form of a chain rule. To make use of this fact, introduce the function $u(t)$ defined as

$$u(t) = f(x(t), y(t)).$$

Note that $u(t)$ is the value of f at a point on the level curve that passes through (a, b). Hence $u(t) = f(a, b)$. What is more important is that $u(t)$ is a constant function. Therefore, $du/dt = 0$.

Now, $u = f(x, y)$, where x and y are functions of t. The chain rule asserts that

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

Since $du/dt = 0$, (16.5.1) follows. Hence ∇f, evaluated at (a, b), is a normal to the level curve of f that passes through (a, b). •

What does this theorem say about the daily weather map that shows the barometric pressure? A level curve, or contour, shows the points where the pressure has a prescribed value. The gradient ∇f at anyplace on such a curve points in the direction in which the pressure increases most rapidly. So $-\nabla f$ points where the pressure is decreasing most rapidly. Since the wind tends to go from high pressure to low pressure, we can think of $-\nabla f$ as representing the wind.

Figure 16.5.1 shows a typical level curve and gradient. The gradient is perpendicular to the level curve. Moreover, as we saw in Section 16.4, the gradient points in the direction in which the function increases most rapidly.

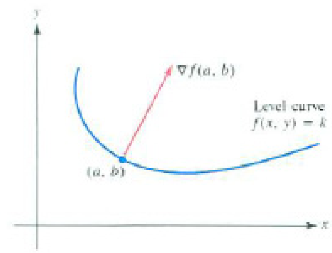


Figure 16.5.1:

EXAMPLE 1 Find and draw a normal vector to the hyperbola $xy = 6$ at the point $(2, 3)$.
SOLUTION Let $f(x, y) = xy$. Then $f_x = y$ and $f_y = x$. Hence,

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

In particular

$$\nabla f(2, 3) = 3\mathbf{i} + 2\mathbf{j}.$$

This gradient and level curve $xy = 6$ are shown in Figure 16.5.2. ◇

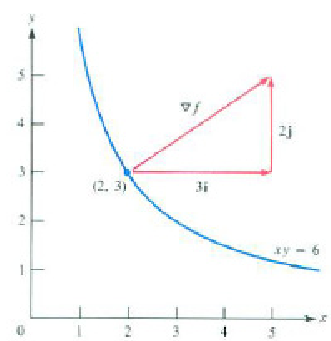


Figure 16.5.2:

EXAMPLE 2 Find an equation of the tangent line to the ellipse $x^2 + 2y^2 = 7$ at the point $(2, 1)$.
SOLUTION As we saw in Section 14.4, we may write the equation of a line in the plane if we know a point on the line and a vector normal to the line. We know that $(2, 1)$ lies on the line. We use a gradient to produce a normal.

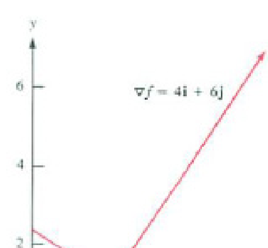
The ellipse $x^2 + 3y^2 = 7$ is a level curve of the function $f(x, y) = x^2 + 3y^2$. Since $f_x = 2x$ and $f_y = 6y$, $\nabla f = 2x\mathbf{i} + 6y\mathbf{j}$. In particular

$$\nabla f(2, 1) = 4\mathbf{i} + 6\mathbf{j}.$$

Hence the tangent line at $(2, 1)$ has an equation

$$4(x - 2) + 6(y - 1) = 0 \qquad \text{or} \qquad 4x + 6y = 14.$$

The level curve, normal vector, and tangent line are all shown in Figure 16.5.3. ◇



Normals to a Surface

We can construct a vector perpendicular to a surface $f(x, y, z) = k$ at a given point $P = (a, b, c)$ as easily as we constructed a vector perpendicular to a planar curve. It turns out that the gradient vector ∇f, evaluated at (a, b, c), is perpendicular to the surface $f(x, y, z) = k$. The proof of this result is similar to the proof for normal vectors to a level curve, given earlier in this section.

Before going on, we must state what is meant by a "vector being perpendicular to a surface."

A vector is perpendicular to a curve at a point (a, b, c) on the curve if the vector is perpendicular to a tangent vector to the curve at (a, b, c).

DEFINITION *(Normal vector to a surface)* A vector is perpendicular to a surface at the point (a, b, c) on this surface if the vector is perpendicular to each curve on the surface through the point (a, b, c). Such a vector is called a **normal vector**.

Finding a Normal to the surface $f(x, y, z) = k$

SHERMAN: You have the note (about using $a\mathbf{i} + b\mathbf{j} + c\mathbf{k}$ or $< a, b, c >$. I think we should use both, but I don't have a strong opinion about this.

Theorem. *Normal vectors to a level surface The gradient ∇f at (a, b, c) is a normal to the level surface of f passing through (a, b, c).*

Proof

Let $\mathbf{G}(t) = x(t)\mathbf{i} + y(t)\mathbf{j} + z(t)\mathbf{k}$ be the parameterizations of a curve in the level surface of f that passes through the point (a, b, c). Assume $\mathbf{G}(t_0) = a\mathbf{i} + b\mathbf{j} + c\mathbf{k}$. Then $\mathbf{G}'(t_0)$ is the tangent vector to the curve at the point (a, b, c) and the gradient at (a, b, c) is

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

We wish to show that

$$\nabla f \cdot \mathbf{G}'(t_0) = 0;$$

that is

$$\frac{\partial f}{\partial x}(a, b, c)x'(t_0) + \frac{\partial f}{\partial y}(a, b, c)y'(t_0) + \frac{\partial f}{\partial z}(a, b, c)z'(t_0) = 0. \tag{16.5.2}$$

Figure 16.5.4:

(See Figure 16.5.4.) Introduce the function $u(t)$ defined by

$$u(t) = f(x(t), y(t), z(t)).$$

Once again the chain rule comes to our aid.

By the chain rule,

$$\left.\frac{du}{dt}\right|_{t=t_0} = \frac{\partial f}{\partial x}(a, b, c)x'(t_0) + \frac{\partial f}{\partial y}(a, b, c)y'(t_0) + \frac{\partial f}{\partial z}(a, b, c)z'(t_0) = 0 \tag{16.5.3}$$

However, *since the curve $\mathbf{G}(t)$ lies on a level surface of f, $u(t)$ is constant.* [In fact, $u(t) = f(a, b, c)$.] Thus $du/dt = 0$, and the right side of (16.5.3) is 0, as required. •

A simple check of this result is to see whether it is correct when the level surfaces are just planes. Consider $f(x, y, z) = Ax + By + Cz + D$. The plane $Ax + By + Cz + D = 0$ is the level surface $f(x, y, z) = 0$. According to the theorem, ∇f is perpendicular to this surface. Now, $f_x = A$, $f_y = B$, and $f_z = C$. Hence,

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

This agrees with the fact that $A\mathbf{i} + B\mathbf{j} + C\mathbf{k} = 0$, as we saw in Section 14.4.

EXAMPLE 3 Find a normal vector to the ellipsoid $x^2 + y^2/4 + z^2/9 = 3$ at the point $(1, 2, 3)$.
SOLUTION The ellipsoid is a level surface of the function

$$f(x, y, z) = x^2 + \frac{y^2}{4} + \frac{z^2}{9}.$$

The gradient of f at the point (x, y, z) is

$$\nabla f = 2x\mathbf{i} + \frac{y}{2}\mathbf{j} + \frac{2z}{9}\mathbf{k}.$$

At $(1, 2, 3)$

$$\nabla f = 2\mathbf{i} + \mathbf{j} + 2/3\mathbf{k}.$$

This vector is normal to the ellipsoid at $(1, 2, 3)$. ◇

Tangent Planes to a Surface

Now that we can find a normal to a surface we can define a tangent plane at a point on the surface.

DEFINITION *(Tangent plane to a surface)* Consider a surface that is a level surface of a function $u = f(x, y, z)$. Let (a, b, c) be a point on this surface where ∇f is not 0. The tangent plane to the surface at the point (a, b, c) is that plane through (a, b, c) that is perpendicular to the vector ∇f evaluated at (a, b, c).

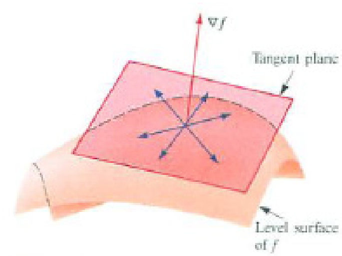


Figure 16.5.5:

The tangent plane at (a, b, c) is the plane that best approximates the surface near (a, b, c). It consists of all the tangent lines at (a, b, c) to curves in the surface that pass through the point (a, b, c). See Figure 16.5.5.

Note that an equation of the tangent plane to the surface $f(x, y, z) = k$ at (a, b, c) is

$$\frac{\partial f}{\partial x}(a, b, c)(x - a) + \frac{\partial f}{\partial y}(a, b, c)(y - b) + \frac{\partial f}{\partial z}(a, b, c)(z - c) = 0.$$

EXAMPLE 4 Find an equation of the tangent plane to the ellipsoid $x^2 + y^2/4 + z^2/9 = 3$ at the point $(1, 2, 3)$.

SOLUTION By Example 3, the vector $2\mathbf{i} + \mathbf{j} + 2/3\mathbf{k}$ is normal to the surface at the point $(1, 2, 3)$. The tangent plane consequently has an equation

$$2(x - 1) + 1(u - 2) + 2/3(z - 3) = 0$$

◇

Normals and Tangent Planes to $z = f(x, y)$

A surface may be described explicitly in the form $z = f(x, y)$ rather than implicitly in the form $f(x, y, z) = k$. The techniques already developed enable us to find the normal and tangent plane in the case $z = f(x, y)$ as well.

Finding a normal to the surface $z = f(x, y)$

We need only rewrite the equation $z = f(x, y)$ in the form $z - f(x, y) = 0$. Then define $g(x, y, z)$ to be $z - f(x, y)$. The surface $z - f(x, y)$ is simply the particular level surface of g given by $g(x, y, z) = 0$. There is no need to memorize an extra formula for a vector normal to the surface $z = f(x, y)$. The next example illustrates this advice.

DOUG: I graphed $z = xy$, not $z = x^2 = y^2$. What to do? SHERMAN: I do not see how this graph is incorrect.

EXAMPLE 5 Find a vector perpendicular to the saddle $z = y^2 - x^2$ at the point $(1, 2, 3)$.

SOLUTION In this case, rewrite $z = y^2 - x^2$ as $z + x^2 - y^2 = 0$. The surface in question is a level surface of $g(x, y, z) = z + x^2 - y^2$. Hence $\nabla g = 2x\mathbf{i} - 2y\mathbf{j} + \mathbf{k}$ is perpendicular to the surface at the point $(1, 2, 3)$.

This surface looks like a saddle near the origin. The surface and the normal vector $2\mathbf{i} - 4\mathbf{j} + \mathbf{k}$ are shown in Figure 16.5.6. ◇

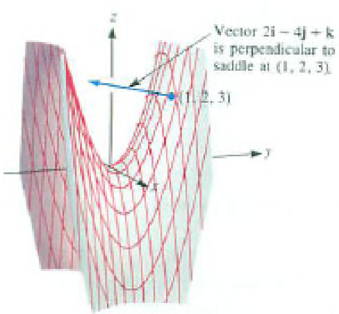


Figure 16.5.6:

Estimates and the Tangent Planes

In the case of a function of one variable, $y = f(x)$, the tangent line at $(a, f(a))$ closely approximates the graph of $y = f(x)$. The equation of the tangent line $y = f(a) + f'(a)(x - a)$ gives us a linear approximation of $f(x)$. (See Section 5.3.)

We can use the tangent plane to the surface $z = f(x, y)$ similarly. To find the equation of the plane tangent at $(a, b, f(a, b))$, we first rewrite the equation of the surface as

$$g(x, y, z) = f(x, y) - z = 0.$$

Then ∇g is a normal to the surface at $(a, b, f(a, b))$. Now,

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

where the partial derivatives are evaluated at (a, b).

The equation of the tangent plane at $(a, b, f(a, b))$ is therefore

$$\frac{\partial f}{\partial x}(a,b)(x-a) + \frac{\partial f}{\partial y}(a,b)(y-b) - (z - f(a,b)) = 0.$$

We can rewrite this equation as

$$z = f(a,b) + \frac{\partial f}{\partial x}(a,b)(x-a) + \frac{\partial f}{\partial y}(a,b)(y-b). \qquad (16.5.4)$$

Letting $\Delta x = x - a$ and $\Delta y = y - b$, (16.5.4) becomes

$$z = \frac{\partial f}{\partial x}(a,b)\Delta x + \frac{\partial f}{\partial y}(a,b)\Delta y.$$

This tells us that the change of the z coordinate on the tangent plane, as the x coordinate changes from a to $a + \Delta x$ and the y coordinate changes from b to $b + \Delta y$ is exactly

$$\frac{\partial f}{\partial x}(a,b)\Delta x + \frac{\partial f}{\partial y}(a,b)\Delta y.$$

By (16.3.1) in Section 16.3, this is an estimate of the change Δf in the function f as its argument changes from (a, b) to $(a + \Delta x, b + \Delta y)$. This is another way of saying that "the tangent plane to the surface $z = f(x, y)$ at $(a, b, f(a, b))$ looks a lot like that surface near that point." See Figure 16.5.7.

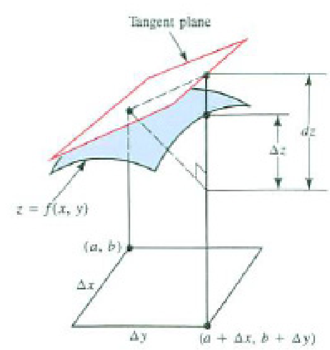


Figure 16.5.7:

The expression $f_x(a, b)\ dx + f_y(a, b)\ dy$ is called the differential of f at (a, b). For small values of dx and dy it is a good estimate of $\Delta f = f(a + dx, b + dy) - f(a, b)$.

EXAMPLE 6 Let $z = f(x, y) = x^2y$. Let $\Delta z = f(1.01, 2.02) - f(1, 2)$ and let

$$dz = \frac{\partial f}{\partial x}(1,2)\cdot 0.0 + \frac{\partial f}{\partial y}(1,2)\cdot 0.02.$$

Compute Δz and dz.

SOLUTION

$$\Delta z = (1.01)^2(2.02) - 1^2 2 = 2.060602 - 2 = 0.060602$$

Since $f_x = 2xy$ and $f_y = x^2$, we have $f_x = 4$ and $f_y = 1$ at $(1, 2)$. Hence,

$$dz = (4)(0.01) + (1)(0.02) = 0.06.$$

Note that dz is a good approximation of Δz. ◇

Function	Level Curve/Surface	Normal	Tangent
$f(x,y)$	level curve: $f(x,y)=k$	$\nabla f = f_x\mathbf{i} + f_y\mathbf{j}$	Tangent line $f_x(a,b)(x-a)+$ $f(a,b)$
$f(x,y,z)$	level surface: $f(x,y,z)=k$	$\nabla f = f_x\mathbf{i} + f_y\mathbf{j} + f_z\mathbf{k}$	Tangent plane $f_x(a,b,c)(x-a)$ $b)+f_z(a,b,c)(z-$

Table 16.5.1: t15-5-1

Summary

This table summarizes mot of what we did concerning normal vectors.

To find a normal and tangent plane to a surface given in the form $z = f(x,y)$, treat the surface as a level surface of the function $z - f(x,y)$, normally $z - f(x,y) = 0$.

We also showed that the differential approximation of Δf in Section 16.3 is simply the change along the tangent plane.

DOUG: Must get implicit diff in partials somewhere?? SHERMAN: Exercises?? Maybe back in the Chain Rule section, with a few more exercises in this section. Or, in §16.8.

EXERCISES for Section 16.5 *Key:* R–routine, M–moderate, C–challenging

1.[R] In estimating the value of a right circular cylindrical tree trunk, a lumber jack may make a 5 percent error in estimating the diameter and a 3 percent error in measuring the height. How large an error may he make in estimating the volume?

2.[R] Let T denote the time it takes for a pendulum to complete a back-and-forth swing. If the length of the pendulum is L and g the acceleration due to gravity, then

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

A 3 percent error may be made in measuring L and a 2 percent error in measuring g. How large an error may we make in estimating T?

3.[R] Let $A(x, y) = xy$ be the area of a rectangle of sides x and y. Compute ΔA and dA and show them in Figure 16.5.8

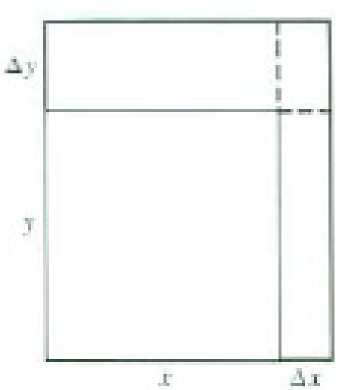


Figure 16.5.8:

The differential of a function $u = f(x, y, z)$ is defined to be $f_x\Delta x + f_y\Delta y + f_z\Delta z$, in analogy with the differential of a function of two variables.

4.[R] Let $V(x, y, z) = xyz$ be the volume of a box of sides x, y, and z. Compute ΔV and dV and show them in Figure 16.5.9.

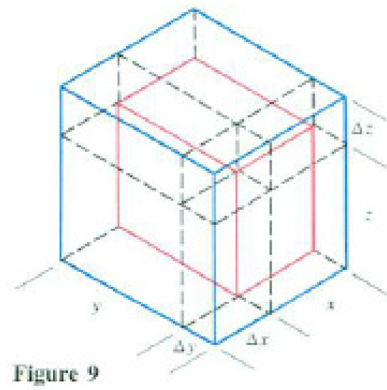


Figure 16.5.9:

5.[R] Let $u = f(x, y, z)$ and $\mathbf{r} = \mathbf{G}(t)$. Then u is a composite function of t. Show that

$$\frac{du}{dt} = \nabla f \cdot \mathbf{G}'(t),$$

where ∇f is evaluated at $\mathbf{G}(t)$. For instance, let $y = f(x, y, z)$ and let $\mathbf{G}$ describe the journey of a bug. Then the rate of change in the temperature as observed by the but is the dot product of the temperature gradient ∇f and the velocity vector $\mathbf{v} = \mathbf{G}'$.

6.[R] We have found a way to find a normal and a tangent plane to a surface. How would you find a *tangent line* to a surface? Illustrate your method by finding a line that is tangent to the surface $z = xy$ at $(2, 3, 6)$.

7.[R] Suppose you are at the point (a, b, c) on the level surface $f(x, y, z) = k$. At that point $\nabla F = 2\mathbf{i} + 3\mathbf{j} - 4\mathbf{k}$.

(a) If $\mathbf{u}$ is tangent to the surface at (a, b, c), what would $D_{\mathbf{u}}f$ equal?

(b) If $\mathbf{u}$ is normal to the level surface at (a, b, c), what would $D_{\mathbf{u}}$ equal? (There are two such normals.)

8.[R]

(a) Draw three level curves of the function f defined by $f(x, y) = xy$. Include the curve through $(1, 1)$ as one of them.

(b) Draw three level curves of the function g defined by $g(x, y) = x^2 - y^2$. Include the curve through $(1, 1)$ as one of them.

(c) Draw three level curves of the function g defined by $g(x, y) = x^2 - y^2$. Include the curve through $(1, 1)$ as one of them.

(d) Prove that each level curve of f intersects each level curve of g at a right angle.

(e) If we think of f as air pressure, how may we interpret the level curves of g?

9.[R]

(a) Draw a level curve for the function $2x^2 + y^2$.

(b) Draw a level curve for the function y^2/x.

(c) Prove that any level curve of $2x^2+y^2$ crosses any level curve of y^2/x at a right angle.

10.[R] The surfaces $x^2yz = 1$ and $xy + yz + zx = 3$ both pass through the point $(1, 1, 1)$. The tangent planes to these surfaces meet in a line. Find parametric equations for this line.

11.[R] Let $T(x, y, z)$ be the temperature at the point (x, y, z), where ∇T is not $\mathbf{0}$. A level surface $T(x, y, z) - k$ is called an *isotherm*. Show that if you are at the point (a, b, c) and wish to move in the direction in which the temperature changes most rapidly, you would move in a direction perpendicular to the isotherm that passes through (a, b, c).

12.[R] Two surfaces $f(x, y, z) = 0$ and $g(x, y, z) = 0$ both pass through the point (a, b, c). Their intersection is a curve. How would you find a tangent vector to that curve at (a, b, c)?

13.[R] Write a short essay on the wonders of the chain rule. Include a description of how it was used to show that $D_{\mathbf{u}}f = \nabla f \cdot \mathbf{u}$ and in showing that ∇f is a normal to the level surface of f at the point where it is evaluated.

The angle between two surfaces that pass through (a, b, c) is defined as the angle between the two lines through (a, b, c) that are perpendicular to the two surfaces at the point (a, b, c). This angle is taken to be acute. Use this definition in Exercises 14 to 16.

14.[R]

(a) Show that the point $(1, 1, 2)$ lies on the surfaces $xyz = 2$ and $x^3yz^2 = 4$.

(b) Find the angle between the surfaces in (a) at the point $(1, 1, 2)$.

15.[R]

(a) Show that the point $(1, 2, 3)$ lies on the plane

$$2x + 3y - z = 5$$

and the sphere

$$x^2 + y^2 + z^2 = 14.$$

(b) Find the angle between them at the point $(1, 2, 3)$.

16.[R]

(a) Show that the surfaces $z = x^2y^3$ and $z = 2xy$ pass through the point $(2, 1, 4)$.

(b) At what angle do they cross at that point?

17.[R] Let $z = f(x, y)$ describe a surface. Assume that at $(3, 5)$, $z = 7$, $\partial z/\partial x = 2$, and $\partial z/\partial y = 3$.

(a) Find two vectors that are tangent to the surface at $(3, 5, 7)$.

(b) Find a normal to the surface at $(3, 5, 7)$.

(c) Estimate $f(3.02, 4.99)$.

18.[R] This map shows the pressure $p(x,y)$ in terms of level curves called isobars. Where is the gradient of p, ∇p the longest? In what direction does it point? In which direction (approximately) would the wind vector point?

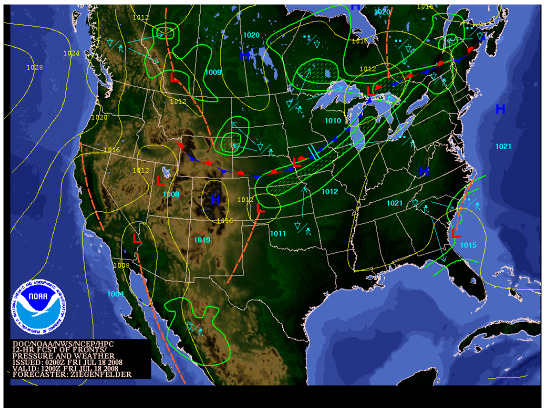


Figure 16.5.10: Source: `http://www.walltechnet.com/b_f/Weather/USAIsobarMap.htm` (18 July 2008)

19.[M] How far is the point $(2,1,3)$ from the tangent plane to $z = xy$ at $(3,4,12)$?

20.[C] The surface $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$ is called an ellipsoid. If $a^2 = b^2 = c^2$ it is a sphere. Show that if a^2, b^2, and c^2 are distinct, then there are exactly six normals on the ellipsis that pass through the origin.

21.[C] Let S be a surface with the property that its target planes are always perpendicular to $\mathbf{r}$. Must S be a sphere?

16.6 Critical Points and Extrema

Recall: $f''(x)$ positive means the graph of f is concave up; $f''(x)$ negative means the graph of f is concave down.

In the case of a function of one variable, $y = f(x)$, the first and second derivatives were of use in searching for relative extrema. First, we looked for critical numbers, that is, solutions of the equation $f'(x) = 0$. Then we checked the value of $f''(x)$ at each such point. If $f''(x)$ were positive, the critical number gave a relative minimum. If $f''(x)$ were negative, the critical number gave a relative maximum. If $f''(x)$ were 0, then anything might happen: a relative minimum or maximum or neither. (For instance, at 0 the functions x^4, $-x^4$, and x^3 have both first and second derivatives equal to 0, but the first function has a relative minimum there, the second has a relative maximum, and the third has neither.) In such a case, we have to resort to other tests.

Remember that $\frac{\partial f}{\partial x} = f_x$. The subscript notation is used in text to save space.

This section extends the idea of a critical point to functions $f(x, y)$ of two variables and shows how to use the second-order partial derivatives f_{xx}, f_{yy}, and f_{xy} to see whether the critical point provides a relative maximum, relative minimum, or neither.

Extrema of $f(x, y)$

The number M is called the **maximum** (or **global maximum**) of f over a set R in the plane if it is the largest value of $f(x, y)$ for (x, y) in R. A **relative maximum** (or **local maximum**) of f occurs at a point (a, b) in R if there is a disk around (a, b) such that $f(a, b) \geq f(x, y)$ for all points (x, y) in the disk. **Minimum** and **relative** (or **local) minimum** are defined similarly.

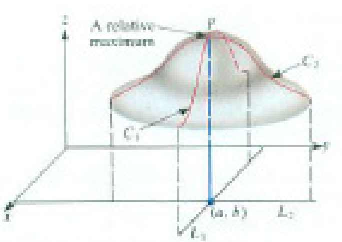


Figure 16.6.1:

Let us look closely at the surface above a point (a, b) where a relative maximum of f occurs. Assume that f is defined for all points within some circle around (a, b) and possesses partial derivatives at (a, b). Let L_1 be the line $y = b$ in the xy plane; let L_2 be the line $x = a$ in the xy plane. (See Figure 16.6.1. Assume, for convenience, that the values of f are positive.)

Let C_1 be the curve in the surface directly above the line L_1. Let C_2 be the curve in the surface directly above the line L_2. Let P be the point on the surface directly above (a, b).

Since f has a relative maximum at (a, b), no point on the surface near P is higher than P. Thus P is a highest point on the curve C_1 and on the curve C_2 (for points near P). The study of functions of one variable showed that both these curves have horizontal tangents at P. In other words, at (a, b) both partial derivatives of f must be 0:

$$\frac{\partial f}{\partial x}(a, b) = 0 \qquad \text{and} \qquad \frac{\partial f}{\partial y}(a, b) = 0.$$

This conclusion is summarized in the following theorem.

Theorem. *Relative Extremum of $f(x, y)$ Let f be defined on a domain that includes the point (a, b) and all points within some circle whose center is (a, b). If f has a relative maximum (or relative minimum) at (a, b) and f_x and f_y exist at (a, b), then both these partial derivatives are 0 at (a, b); that is,*

$$\frac{\partial f}{\partial x}(a, b) = 0 = \frac{\partial f}{\partial y}(a, b),$$

In short, the gradient of f, ∇f is $\mathbf{0}$ *at a relative extremum.*

A point (a, b) where both partial derivatives f_x and f_y are 0 is clearly of importance. The following definition is analogous to that of a critical point of a function of one variable.

> **DEFINITION** *(Critical point)* If $f_x(a, b) = 0$ and $f_y(a, b) = 0$, the point (a, b) is a **critical point** of the function $f(x, y)$.

You might expect that if (a, b) is a critical point of f and the two second partial derivatives f_{xx} and f_{yy} are both positive at (a, b), then necessarily has a relative minimum at (a, b). The next example shows that *the situation is not that simple.*

EXAMPLE 1 Find the critical points of $f(x, y) = x^2 + 3xy + y^2$ and determine whether there is an extremum there.

SOLUTION First, find any critical points by setting both f_x and f_y equal to 0. This gives the simultaneous equations

$$2x + 3y = 0 \qquad \text{and} \qquad 3x + 2y = 0.$$

Since the only solution of these equations is $(x, y) = (0, 0)$, the function has one critical point, namely $(0, 0)$.

Now look at the graph of f for (x, y) near $(0, 0)$.

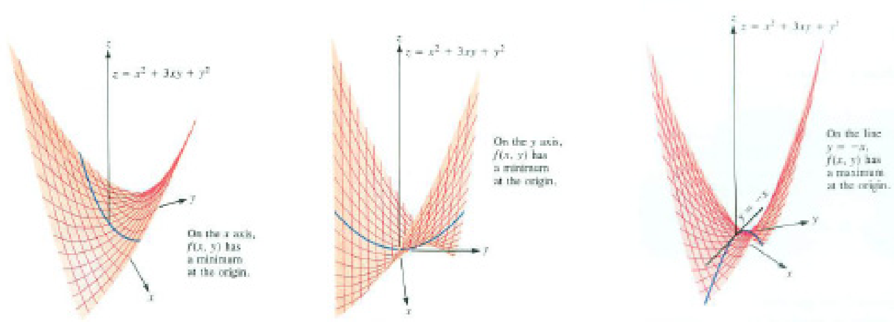
First, consider how f behaves for points on the x axis. We have $f(x, 0) = x^2 + 3 \cdot x \cdot 0 + 0^2 = x^2$. Therefore, considered *only as a function of* x, the function has a minimum at the origin. (See Figure 16.6.2(a).)

On the y-axis, the function reduces to $f(0, y) = y^2$, whose graph is another parabola with a minimum at the origin. (See Figure 16.6.2(b).) Note also that $f_{xx} = 2$ and $f_{yy} = 2$, so both are positive at $(0, 0)$.

So far, the evidence suggests that f has a relative minimum at $(0, 0)$. However, consider its behavior on the line $y = -x$. For points (x, y) on this line

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

On this line the function assumes negative values, and its graph is a parabola opening downward, as shown in Figure 16.6.2(c).



(a) (b) (c)

Figure 16.6.2:

Thus $f(x, y)$ has neither a relative maximum nor minimum at the origin. Its graph resembles a saddle. ◇

f_{xx} and f_{yy} describe the behavior of $f(x, y)$ only on lines parallel to the x-axis and y-axis, respectively.

Example 1 shows that to determine whether a critical point of $f(x, y)$ provides an extremum, it is not enough to look at f_{xx} and f_{yy} The criteria are more complicated and involve the mixed partial derivative f_{xy} as well. Exercise 58 outlines a proof of the following theorem. At the end of this section a proof is presented in the special case when $f(x, y)$ is a polynomial of the form $Ax^2 + Bxy + Cy^2$, where A, B and C are constants.

In subscript notation, $D = f_{xx}f_{yy} - (f_{xy})^2$.

Theorem 16.6.1. *Second-partial-derivative test for $f(x, y)$ Let (a, b) be a critical point of the function $f(x, y)$. Assume that the partial derivatives f_x, f_y, f_{xx}, f_{xy}, and f_{yy} are continuous at and near (a, b). Let*

$$D = \frac{\partial^2 f}{\partial x^2}(a, b)\frac{\partial^2 f}{\partial y^2}(a, b) - \left(\frac{\partial^2 f}{\partial x \partial y}(a, b)\right)^2.$$

1. *If $D > 0$ and $f_{xx}(a, b) > 0$, then f has a relative minimum at (a, b).*

2. *If $D > 0$ and $f_{xx}(a, b) < 0$, then f has a relative maximum at (a, b).*

3. *If $D < 0$, then f has neither a relative minimum nor a relative maximum at (a, b). (There is a* ***saddle point*** *at (a, b).)*

If $D = 0$, then anything can happen; there may be a relative minimum, a relative maximum, or a saddle. These possibilities are illustrated in Exercise 43.

To see what the theorem says, consider case 1, the test for a relative minimum. It says that $f_{xx}(a, b) > 0$) (which is to be expected) and that

$$\frac{\partial^2 f}{\partial x^2}(a, b)\frac{\partial^2 f}{\partial y^2}(a, b) - \left(\frac{\partial^2 f}{\partial x \partial y}(a, b)\right)^2 > 0,$$

Or equivalently,

$$\left(\frac{\partial^2 f}{\partial x \partial y}(a,b)\right)^2 < \frac{\partial^2 f}{\partial x^2}(a,b)\frac{\partial^2 f}{\partial y^2}(a,b). \qquad (16.6.1)$$

Memory aid regarding size of f_{xy}

Since the square of a real number is never negative, and $f_{xx}(a,b)$ is positive, it follows that $f_{yy}(a,b) > 0$, which was to be expected. But inequality (16.6.1) says more. It says that the *mixed partial $f_{xy}(a,b)$ must not be too large.* For a relative maximum or minimum, inequality (16.6.1) must hold. This may be easier to remember than "$D > 0$."

EXAMPLE 2 Examine each of these functions for relative extrema:

1. $f(x,y) = x^2 + 3xy + y^2$,
2. $g(x,y) = x^2 + 2xy + y^2$,
3. $h(x,y) = x^2 + xy + y^2$.

SOLUTION

1. The case $f(x,y) = x^2 + 3xy + y^2$ is Example 1. The origin is the only critical point, and it provides neither a relative maximum nor a relative minimum. We can check this by the use of the discriminant. We have

$$\frac{\partial^2 f}{\partial x^2}(0,0) = 2, \qquad \frac{\partial^2 f}{\partial x \partial y}(0,0) = 3, \qquad \text{and} \qquad \frac{\partial^2 f}{\partial y^2}(0,0) = 2.$$

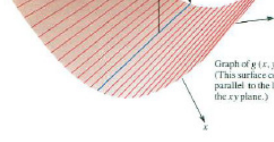
Hence $D = 2 \cdot 2 - 3^2 = -5$ is negative. By the second-partial-derivative test, there is neither a relative maximum nor a relative minimum at the origin. Instead, there is a saddle there.

2. It is straightforward matter to show that all the points on the line $x+y = 0$ are critical points of $g(x,y) = x^2 + 2xy + y^2$. Moreover,

$$\frac{\partial^2 g}{\partial x^2}(x,y) = 2, \qquad \frac{\partial^2 g}{\partial x \partial y}(x,y) = 2, \qquad \text{and} \qquad \frac{\partial^2 g}{\partial y^2}(x,y) = 2.$$

Thus the discriminant $D = 2 \cdot 2 - 2^2 = 0$. Since $D = 0$, the discriminant gives no information.

Note, however, that $x^2 + 2xy + y^2 = (x+y)^2$ and so, being the square of a real number, is always greater than or equal to 0. Hence the origin provides a relative minimum of $x^2 + 2xy + y^2$. (In fact, any point on the line $x + y = 0$ provides a relative minimum. Since $g(x,y) = (x+y)^2$, the function is constant on each line $x + y = c$, for any choice of the constant c. See Figure 16.6.3.)



Graph of $g(x,y) = x^2 + 2xy + y^2$ (This surface consists of lines parallel to the line $x + y = 0$ in the xy plane.)

3. For $h(x, y) = x^2 + xy + y^2$, again the origin is the only critical point and we have

$$\frac{\partial^2 h}{\partial x^2}(0,0) = 2, \qquad \frac{\partial^2 h}{\partial x \partial y}(0,0) = 1, \qquad \text{and} \qquad \frac{\partial^2 h}{\partial y^2}(0,0) = 2.$$

In this case, $D = 2 \cdot 2 - 1^2 = 3$ is positive and $h_{xx}(0,0) > 0$. Hence $x^2 + xy + y^2$ has a relative minimum at the origin.

The graph of h is shown in Figure 16.6.4

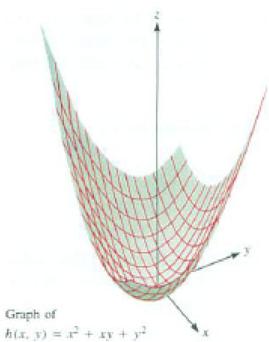


Figure 16.6.4: Function has no global extrema.

◇

EXAMPLE 3 Examine $f(x, y) = x + y + 1/(xy)$ for global and relative extrema.

SOLUTION When x and y are both large positive numbers or small positive numbers, then $F(x, y)$ may be arbitrarily large. There is therefore no global maximum. By allowing x and y to be negative numbers of large absolute values, we see that there is no global minimum.

Any local extrema will occur at a critical point. We have

$$\frac{\partial f}{\partial x} = 1 - \frac{1}{x^2 y} \qquad \text{and} \qquad \frac{\partial f}{\partial y} = 1 - \frac{1}{xy^2}.$$

Setting these derivatives equal to 0 gives

$$\frac{1}{x^2 y} = 1 \qquad \text{and} \qquad \frac{1}{xy^2} = 1 \tag{16.6.2}$$

Hence $x^2 y = xy^2$. Since the function f is not defined when x or y is 0, we may assume $xy \neq 0$. Dividing both sides of $x^2 y = xy^2$ by xy gives $x = y$. By (16.6.2) (either equation), $1/x^3 = 1$; hence $x = 1$. Thus there is only one critical point, namely, $(1, 1)$.

To find whether it is a relative extremum, use Theorem 16.6.1. We have

$$\frac{\partial^2 f}{\partial x^2} = \frac{2}{x^3 y}, \qquad \frac{\partial^2 f}{\partial x \partial y} = \frac{1}{x^2 y^2}, \qquad \text{and} \qquad \frac{\partial^2 f}{\partial y^2} = \frac{2}{xy^3}.$$

Thus at $(1, 1)$,

$$\frac{\partial^2 f}{\partial x^2} = 2, \qquad \frac{\partial^2 f}{\partial x \partial y} = 1, \qquad \text{and} \qquad \frac{\partial^2 f}{\partial y^2} = 2.$$

Therefore,

$$D = 2 \cdot 2 - 1^2 = 3 > 0.$$

Since $D > 0$ and $f_{xx})(1, 1) > 0$, the point $(1, 1)$ provides a relative minimum. ◇

Extrema on a Bounded Region

In Section 4.3, we saw how to find a maximum of a differentiable function, $y = f(x)$, on an interval $[a, b]$. The procedure is as follows:

1. First find any numbers x in $[a, b]$ (other than a or b) where $f'(x) = 0$. Such a number is called a critical number. If there are no critical numbers, the maximum occurs at a or b.

2. If there are critical numbers, evaluate f at them. Also find the values of $f(a)$ and $f(b)$. The maximum of f in $[a, b]$ is the largest of the numbers: $f(a)$, $f(b)$, and the values of f at critical numbers.

We can similarly find the maximum of $F(x, y)$ in a region R in the plane bounded by some polygon or curve. (See Figure 16.5.7.) It is assumed that R includes its border and is a finite region in the sense that it lies within some disk. (In advanced calculus, it is proved that a continuous function defined on such a domain has a maximum – and a minimum – value.) If f has continuous partial derivatives, the procedure for finding a maximum is similar to that for maximizing a function on a closed interval.

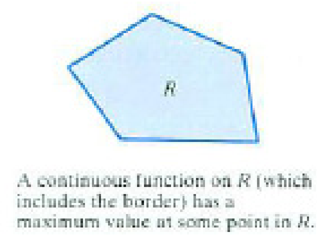


Figure 16.6.5:

1. First find any points that are in R but not on the boundary of R where both f_x and f_y are 0. These are called **critical points**. (if there are no critical points, the maximum occurs on the boundary.)

2. If there are critical points, evaluate f at them. Also find the maximum of f on the boundary. The maximum of f on R is the largest value of f on the boundary and at critical points.

A similar procedure finds the minimum value on a bounded region.

EXAMPLE 4 Maximize the function $f(x, y) = xy(108-2x-2y) = 108xy-2x^2y - 2xy^2$ on the triangle R bounded by the x-axis, the y-axis, and the line $x + y = 54$. (See Figure 16.6.6.)

SOLUTION First find any critical points. We have

$$\frac{\partial f}{\partial x}(x, y) = 108y - 4xy - 2y^2 = 0 \quad (16.6.3)$$

$$\frac{\partial f}{\partial y}(x, y) = 108x - 2x^2 - 4xy = 0 \quad (16.6.4)$$

Find all critical points.

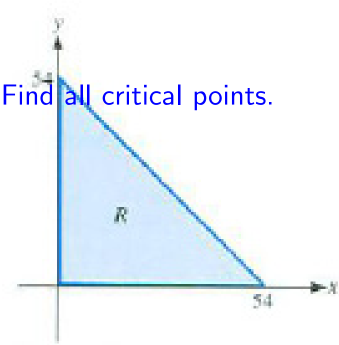


Figure 16.6.6:

which give the simultaneous equations

$$2y(54 - 2x - y) = 0, \quad (16.6.5)$$

$$2x(54 - x - 2y) = 0. \quad (16.6.6)$$

By the first equation, $y = 54-2x$. Substitution of this into the second equation gives: $54 - x - 2(54 - 2x) = 0$, or $-54 + 3x = 0$. Hence $x = 18$ and therefore $y = 54 - 2 \cdot 18 = 18$.

Evaluate f at critical points.

The point $(18, 18)$ lies in the interior of R, since it lies above the x-axis, to the right of the y-axis, and below the line $x + y = 54$. Furthermore, $f(18, 18) = 18 \cdot 18(108 - 2 \cdot 18 - 2 \cdot 18) = 11,664$.

Evaluate f on boundary.

Next we examine the function $f(x, y) = xy(108-2x-2y)$ on the boundary of the triangle R. On the base of R, $y = 0$, so $f(x, y) = 0$. On the left edge of R, $x = 0$, so again $f(x, y) = 0$. On the slanted edge, which lies on the same line $x + y = 54$, we have $108 - 2x - 2y = 0$, so $f(x, y) = 0$ on this edge also. Thus $f(x, y) = 0$ on the entire boundary.

Therefore, the local maximum occurs at the critical point $(18, 18)$ and has the value $11,664$. ◇

EXAMPLE 5 The combined length and girth (distance around) of a package sent through the mail cannot exceed 108 inches. If the package is a rectangle box, how large can its volume be?

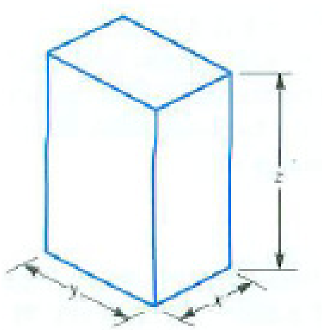


Figure 16.6.7:

SOLUTION Introduce letters to name the quantities of interest. We label its length (a longest side) z and the other sides x and y, as in Figure 16.6.7. The volume $V = xyz$ is to be maximized, subject to girth plus length at most 108, that is,

$$2x + 2y + z \le 108.$$

Since we want the largest box, we might as well restrict our attention to boxes for which

$$2x + 2y + z = 108. \qquad (16.6.7)$$

By (16.6.7), $z = 108 - 2x - 2y$. Thus $V = xyz$ can be expressed as a function of two variables:

$$V = f(x, y) = xy(108 - 2x - 2y).$$

Why is $2x + 2y \le 108$?

This function is to be maximized on the triangle described by $x \ge 0$, $y \ge 0$, $2x + 2y \le 108$, that is, $x + y \le 54$.

These are the same function and region as in the previous example. Hence, the largest box has $x = y = 18$ and $z = 108-2x-2y = 108-2\cdot18-2\cdot18 = 36$; its dimensions are 18 inches by 18 inches by 36 inches and its volume is $11,664$ cubic inches. ◇

Remark: In Example 5 we let z be the length of a longest side, an assumption that was never used. So if the Postal Service regulations read "The length of one edge plus the girth around the other edges shall not exceed 108 inches," the effect would be the same. You would not be able to send a larger box by, say, measuring the girth around the base formed by its largest edges.

EXAMPLE 6 Find the maximum and minimum values of $f(x, y) = x^2 + y^2 - 2x - 4y$ on the disk R of radius 3 and center $(0, 0)$.

SOLUTION First, find any critical points. We have

$$\frac{\partial f}{\partial x} = 2x - 2 \qquad \text{and} \qquad \frac{\partial f}{\partial y} = 2y = 4.$$

The equations

$$\begin{aligned} 2x - 2 &= 0 \\ 2y - 4 &= 0 \end{aligned}$$

have the solutions $x = 1$ and $y - 2$. This point lies in R (since its distance from the origin is $\sqrt{1^2 + 2^2} = \sqrt{5}$, which is less than 3). At the critical point $(1, 2)$, the value of the function is $1^2 + 2^2 - 2(1) - 4(2) = 5 - 2 - 8 = -5$.

Second, find the behavior of f on the boundary, which is a circle of radius 3. We parameterize this circle:

$$\begin{aligned} x &= 3\cos(\theta) \\ y &= 3\sin(\theta). \end{aligned}$$

On this circle,

$$\begin{aligned} f(x, y) &= x^2 + y^2 - 2x - 4y \\ &= (3\cos(\theta))^2 + (3\sin(\theta))^2 - 2(3\cos(\theta)) - 4(3\sin(\theta)) \\ &= 9\cos^2(\theta) + 9\sin^2(\theta) - 6\cos(\theta) - 12\sin(\theta) \\ &= 9 - 6\cos(\theta) - 12\sin(\theta). \end{aligned}$$

We now must find the maximum and minimum of the single-variable function $g(\theta) = 9 - 6\cos(\theta) - 12\sin(\theta)$ for θ in $[0, 2\pi]$.

To do this, find $g'(\theta)$:

$$g'(\theta) = 6\sin\theta - 12\cos\theta.$$

Setting $g'(\theta) = 0$ gives

$$0 = 6\sin(\theta) - 12\cos(\theta)$$

or

$$\sin(\theta) = 2\cos(\theta). \qquad (16.6.8)$$

Why is $\cos(\theta)$ not 0?

To solve (16.6.8), divide by $\cos(\theta)$ (which will not be 0), getting

$$\frac{\sin(\theta)}{\cos(\theta)} = 2$$

or

$$\tan(\theta) = 2.$$

There are two angles θ in $[0, 2\pi]$ such that $\tan(\theta) = 2$. One is in the first quadrant, $\theta = \arctan(2)$, and the other is in the third quadrant, $\pi + \arctan(2)$. To evaluate $g(\theta) = 9 - 6\cos(\theta) - 12\sin(\theta)$ at these angles, we must compute their cosine and sine. The right triangle in Figure 16.6.8 helps us do this.

Figure 16.6.8:

Inspection of Figure 16.6.8 shows that for $\theta = \arctan(2)$,

$$\cos(\theta) = \frac{1}{\sqrt{5}} \qquad \text{and} \qquad \sin(\theta) = \frac{2}{\sqrt{5}}.$$

For this angle

$$g(\arctan(2)) \;=\; 9 - 6\left(\frac{1}{\sqrt{5}}\right) - 12\left(\frac{2}{\sqrt{5}}\right) = 9 - \frac{30}{\sqrt{5}} \approx -4.41641.$$

When $\theta = \pi + \arctan(2)$,

$$\cos(\theta) = \frac{-1}{\sqrt{5}} \qquad \text{and} \qquad \sin(\theta) = \frac{-2}{\sqrt{5}}.$$

So

$$g(\pi + \arctan(2)) \;=\; 9 - 6\left(\frac{-1}{\sqrt{5}}\right) - 12\left(\frac{-2}{\sqrt{5}}\right)$$
$$= 9 + \frac{30}{\sqrt{5}} \approx 22.41641.$$

Since $g(2\pi) = g(0) = 9 - 6(1) - 12(0) = 3$, the maximum of f on the border of R is about 22.41641 and the minimum is about -4.41641. (Recall that at the critical point the value of f is -5.)

We conclude that the maximum value of f on R is about 22.41641 and the minimum value is -5 (and it occurs at the point $(1, 2)$, which is not on the boundar)]. See Figure 16.6.9. ◇

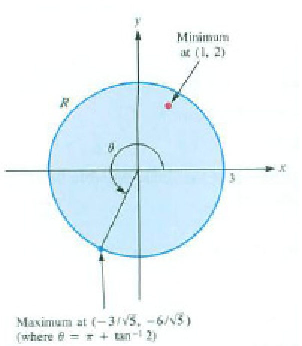


Figure 16.6.9:

Proof of Theorem 16.6.1 in a Special Case

We will prove Theorem 16.6.1 in case $f(x, y)$ is a second-degree polynomial of the form

$$f(x, y) = Ax^2 + Bxy + Cy^2.$$

Theorem 16.6.2. *Let $f(x, y) = Ax^2 + Bxy + Cy^2$, where A, B, and C are constants. Then $(0, 0)$ is a critical point. Let*

$$D = \frac{\partial^2 f}{\partial x^2}(0,0)\frac{\partial^2 f}{\partial y^2}(0,0) - \left(\frac{\partial^2 f}{\partial x \partial y}(0,0)\right)^2.$$

1. *If $D > 0$ and $f_{xx}(0,0) > 0$, then f has a relative minimum at $(0,0)$.*
2. *If $D > 0$ and $f_{xx}(0,0) < 0$, then f has a relative maximum at $(0,0)$.*
3. *If $D < 0$, then f has neither a relative minimum nor a relative maximum at $(0,0)$.*

Proof

We prove Case 1, leaving Cases 2 and 3 for Exercises 60 and 61.

First, compute the first- and second-order partial derivatives of f:

$$\frac{\partial f}{\partial x} = 2Ax + By, \qquad \frac{\partial f}{\partial y} = Bx + 2Cy,$$

$$\frac{\partial^2 f}{\partial x^2} = 2A, \quad \frac{\partial^2 f}{\partial x \partial y} = B, \quad \frac{\partial^2 f}{\partial y^2} = 2C.$$

Note that both f_x and f_y are 0 at $(0,0)$. Hence $(0,0)$ is a critical point and $f(0,0) = 0$. We must show that $f(x,y) \geq 0$ for (x,y) near $(0,0)$. [In fact we will show that $f(x,y) \geq 0$ for all (x,y).]

Next, expressing Case 1 in terms of A, B, and C, we have

$$D = f_{xx}(0,0)f_{yy}(0,0) - f_{xy}^2(0,0) = (2A)(2C) - B^2 = 4AC - B^2 > 0.$$

and $f_{xx}(0,0) = 2A > 0$. In short, we are assuming that $4AC - B^2 > 0$ and $A > 0$, and want to deduce that $f(x,y) = Ax^2 + Bxy + Cy^2 \geq 0$, for (x,y) near $(0,0)$.

Since A is positive, this amounts to showing that

$$A(Ax^2 + Bxy + Cy^2) \geq 0. \tag{16.6.9}$$

We multiply by A to simplify completing the square on the next step.

Now we complete the square,

$$\begin{aligned} A(Ax^2 + Bxy + Cy^2) &= A^2x^2 + ABxy + ACy^2 \\ &= A^2x^2 + ABxy + \frac{B^2}{4}y^2 - \frac{B^2}{4}y^2 + ACy^2 \\ &= (Ax + \frac{B}{2}y)^2 + (AC - \frac{B^2}{4})y^2 \\ &= (Ax + \frac{B}{2}y)^2 + (\frac{4AC - B^2}{4})y^2. \end{aligned}$$

Now, $\left(Ax + \frac{B}{2}y\right)^2 \geq 0$ and $y^2 \geq 0$ since they are squares of real numbers. But by our assumption on D, $4AC - B^2$ is positive. Thus (16.6.9) holds for all (x,y), not just for (x,y) near $(0,0)$ varies Case 1 of the theorem is proved. •

Summary

We defined a critical point of $f(x, y)$ as a point where both partial derivatives f_x and f_y are 0. Even if f_{xx} and f_{yy} are negative there, such a point need not provide a relative maximum. We must also know that f_{xy} is not too large in absolute value.

- If $f_{xx} < 0$ and $f_{xy}^2 < f_{xx}f_{yy}$, then there is indeed a relative maximum at the critical point. (Note that the two inequalities imply $f_{yy} < 0$.)
- Similar criteria hold for a relative maximum: if $f_{xx} > 0$ and $f_{xy}^2 < f_{xx}F_{yy}$, then this critical point is a relative minimum.
- The critical point is a saddle point when $f_{xy} > f_{xx}f_{yy}$.
- When $f_{xy}^2 = f_{xx}f_{yy}$, the critical point may be a relative maximum, relative minimum, or neither.

We also saw how to find extrema of a function defined on a bounded region.

EXERCISES for Section 16.6 *Key:* R–routine, M–moderate, C–challenging

Use Theorems 16.6 and 16.6.1 to determine any relative maxima or minima of the functions in Exercises 1 to 10.

1.[R] $x^2 + 3xy + y^2$

2.[R] $f(x, y) = x^2 - y^2$

3.[R] $f(x, y) = x^2 - 2xy + 2y^2 + 4x$

4.[R] $f(x, y) = x^4 + 8x^2 + y^2 - 4y$

5.[R] $f(x, y) = x^2 - xy + y^2$

6.[R] $f(x, y) = x^2 + 2xy + 2y^2 + 4x$

7.[R] $f(x, y) = 2x^2 + 2xy + 5y^2 + 4x$

8.[R] $f(x, y) = -4x^2 - xy - 3y^2$

9.[R] $f(x, y) = 4/x + 2/y + xy$

10.[R] $f(x, y) = x^3 - y^3 + 3xy$

Let f by a function of x and y such that at (a, b) both f_x and f_y equal 0. In each of Exercises 11 to 16. values are specified for f_{xx}, f_{xy}, and f_{yy} at (a, b). Assume that all these partial derivatives are continuous. On the basis of the given information decides whether

1. f has a relative maximum at (a, b),
2. f has a relative minimum at (a, b),
3. f has a saddle point at (a, b),
4. there is inadequate information.

11.[R] $f_{xy} = 4$, $f_{xx} = 2$, $f_{yy} = 8$

12.[R] $f_{xy} = -3$, $f_{xx} = 2$, $f_{yy} = 4$

13.[R] $f_{xy} = 3$, $f_{xx} = 2$, $f_{yy} = 4$

14.[R] $f_{xy} = 2$, $f_{xx} = 3$, $f_{yy} = 4$

15.[R] $f_{xy} = -2$, $f_{xx} = -3$, $f_{yy} = -4$

16.[R] $f_{xy} = -2$, $f_{xx} = 3$, $f_{yy} = -4$

In Exercises 17 to 24 find the critical points and the relative extrema of the given functions.

17.[R] $x + y - \frac{1}{xy}$

18.[R] $3xy - x^3 - y^3$

19.[R] $12xy - x^3 - y^3$

20.[R] $6xy - x^2y - xy^2$

21.[R] $\exp(x^3 + y^3)$

22.[R] 2^{xy}

23.[R] $3x + xy + x^2y - 2y$

24.[R] $x + y + \frac{8}{xy}$

DH: Are there some more varied realistic exercises? SHERMAN: Has anyone had complaints about these problems in your earlier books? We can look, but I don't believe it's too critical to be creative here.

25.[R] Find the dimensions of the open rectangular box of volume 1 of smallest surface area. Use Theorem 16.6.1 as a check that the critical point provides a minimum.

26.[R] The material for the top and bottom of a rectangular box costs 3 cents per square foot, and that for the sides 2 cents per square foot. What is the least expensive box that has a volume of 1 cubic foot? Use Theorem 16.6.1 as a check that the critical point provides a minimum.

27.[R] UPS ships packages whose combined length and girth is at most 165 inches (and weighs at most 150 pounds).

(a) What are the dimensions of the package with the largest volume that it ships?

(b) What are the dimensions of the package with maximum surface are that UPS will ship?

28.[R] Let $P_1 = (x_1, y_1)$, $P_2 = (x_2, y_2)$, $P_3 = (x_3, y_3)$, and $P_4 = (x_4, y_4)$. Find the coordinates of the point P that minimizes the sum of the squares of the distances from P to the four points.

29.[R] Find the dimensions of the rectangle box of largest volume if its total surface area is to be 12 square meters.

30.[R] Three nonnegative numbers x, y, and z have the sum 1.

(a) How small can $x^2 + y^2 + z^2$ be?

(b) How large can it be?

31.[R] Each year a firm can produce r radios and t television sets at a cost of $2r^2+rt+2t^2$ dollars. It sells a radio for \$600 and a television set for \$900.

(a) What is the profit from the sale of r radios and t television sets? NOTE: Profit is revenue less the cost.

(b) Find the combination of r and t that maximizes profit. Use the discriminant as a check.

32.[R] Find the dimensions of the rectangular box of largest volume that can be inscribed in a sphere of radius 1.

33.[R] For which values of the constant k does $x^2 + kxy + 3y^2$ have a relative minimum at $(0,0)$?

34.[R] For which values of the constant k does the function $kx^2 + 5xy + 4y^2$ have a relative minimum at $(0,0)$?

35.[R] Let $f(x,y) = (2x^2 + y^2)e^{-x^2-y^2}$.

(a) Find all critical points of f.

(b) Examine the behavior of f when x^2+y^2 is large.

(c) What is the minimum value of f?

(d) What is the maximum value of f?

36.[R] Find the maximum and minimum values of the function in Exercise 35 on the circle

(a) $x^2 + y^2 = 1$,

(b) $x^2 + y^2 = 4$.

HINT: Express the function in terms of θ.

37.[R] Find the maximum value of $f(x,y) = 3x^2 - 4y^2 + 2xy$ for points (x,y) in the square region whose vertices are $(0,0)$, $(0,1)$, $(1,0)$, and $(1,1)$.

38.[R] Find the maximum value of $f(x,y) = xy$ for points (x,y) in the triangular region whose vertices are $(0,0)$, $(1,0)$, and $(0,1)$.

39.[R] Maximize the function $-x + 3y + 6$ on the quadrilateral whose vertices are $(1,1)$, $(4,2)$, $(0,3)$, and $(5,6)$.

40.[M]

(a) Show that $z = x^2 - y^2 + 2xy$ has no maximum and no minimum.

(b) Find the minimum and maximum of z if we consider only (x,y) on the circle of radius 1 and center $(0,0)$, that is all (x,y) such that $x^2+y^2 = 1$.

(c) Find the minimum and maximum of z if we consider all (x,y) in the disk of radius 1 and center $(0,0)$, that is, all (x,y) such that $x^2 + y^2 \leq 1$.

41.[M] Suppose z is a function of x and y with continuous second partial derivatives. If, at the point (x_0, y_0), $z_x = 0 = z_y$, $z_{xx} = 3$, and $z_{yy} = 12$, for what values of z_{xy} is it certain that z has a relative minimum at (x_0, y_0)?

42.[M] Let $U(x,y,z) = x^{1/2}y^{1/3}z^{1/6}$ be the "utility" or "desirability" to a given consumer of the amounts x, y, and z of three different commodities. Their prices are, respectively, 2 dollars, 1 dollar, and 5 dollars, and the consumer has 60 dollars to spend. How much of each product should he buy to maximize the utility?

43.[M] This exercise shows that if the discriminant D is 0, then any of the three outcomes mentioned in Theorem 16.6.1 are possible.

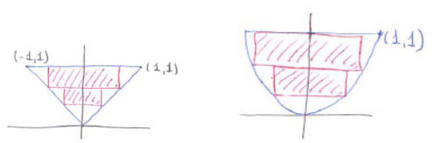
(a) Let $f(x,y) = x^2 + 2xy + y^2$. Show that at $(0,0)$ both f_x and f_y are 0, f_{xx} and f_{yy} are positive, $D = 0$, and f has a relative minimum.

(b) Let $f(x,y) = x^2 + 2xy + y^2 - x^4$ Show that at $(0,0)$ both f_x and f_y are 0, f_{xx} and f_{yy} are positive, $D = 0$, and f has neither a relative maximum nor a relative minimum at $(0,0)$.

(c) Give an example of a function $f(x,y)$ for which $(0,0$ is a critical point and $D = 0$ there, but f has a relative maximum at $(0,0)$.

44.[M] Let $f(x,y) = ax + by + c$, for constants a, b, and c. Let R be a polygon in the xy plane. Show that the maximum and minimum values of $f(x,y)$ on R are assumed only at vertices of the polygon.

45.[M] Two rectangles are placed in the triangle whose vertices are $(0,0)$, $(1,1)$, and $(-1,1)$ as shown in Figure 16.6.10(a).



(a) (b)

Figure 16.6.10:

Show that they can fill as much as 2/3 of the area of the triangle.

46.[M] Two rectangles are placed in the parabola $y = x^2$ as shown in Figure 16.6.10(b). How large can their total area be?

47.[M] Let $P_0 = (a,b,c)$ be a point not on the surface $f(x,y,z) = 0$. Let P be the point on the surface nearest P_0. Show that $\overrightarrow{PP_0}$ is perpendicular to the surface at P. Hint: Show it is perpendicular to each curve on the surface that passes through P.

48.[C] Let (x_1,y_1), (x_2,y_2), ..., (x_n,y_n) be n points in the plane. Statisticians define the **line of regression** as the line that minimizes the sum of the squares of the differences between y_i and the ordinates of the line at x_i. (See Figure 16.6.11.) Let the typical line in the plane have the equation $y = mx + b$.

(a) Show that the line of regression minimizes the sum $\sum_{i=1}^n (y_i - (mx_i + b))^2$ considered as a function of m and b.

(b) Let $f(m,b) = \sum_{i=1}^n (y_i - (mx_i + b))^2$. Compute f_m and f_b.

(c) Show that when $f_m = 0 = f_b b$, we have

$$m\sum_{i=1}^n x_i^2 + b\sum_{i=1}^n x_1 = \sum_{i=1}^n x_i y_i$$

and

$$m\sum_{i=1}^n x_i + nb = \sum_{i=1}^n y_i.$$

(d) When do the simultaneous equations in (c) have a unique solution for m and b?

SHERMAN:I modified your picture in (b), some. OK? I didn't think it was so bad; see my answer.

(e) Find the regression line for the points $(1,1)$, $(2,3)$, and $(3,5)$.

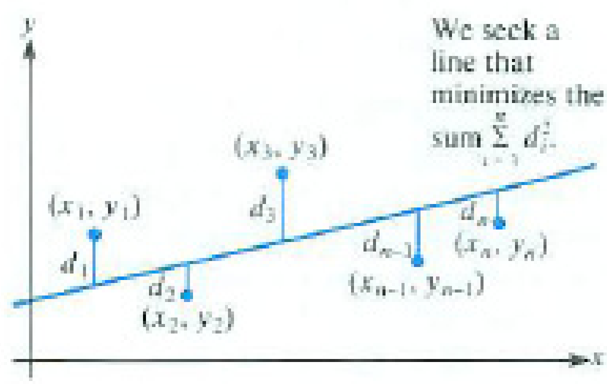


Figure 16.6.11:

49.[C] If your calculator is programmed to compute lines of regression, find and draw the line of regression for the points $(1,1)$, $(2,1.5)$, $(3,3)$, $(4,2)$ and $(5,3.5)$.

50.[C] Let $f(x,y) = (y - x^2)(y - 2x^2)$.

(a) Show that f has neither a local minimum nor a local maximum at $(0,0)$

(b) Show that f has a local minimum at $(0,0)$ when considered only on any fixed line through $(0,0)$.

Suggestion for (b): Graph $y = x^2$ and $y = 2x^2$ and show where $f(x,y)$ is positive and where it is negative.

51.[C] Find (a) the minimum value of xyz, and (b) the maximum value of xyz, for all triplets of nonnegative real numbers x, y, z such that $x + y + z = 1$.

52.[C]

(a) Deduce from Exercise 51 that for any three nonnegative numbers a, b, and c, $\sqrt[3]{abc} \leq (a + b + c)/3$. NOTE: This exercise asserts that the "geometric mean" of three numbers is not larger than their 'arithmetic mean".

(b) Obtain a corresponding result for four numbers.

53.[C] Prove case 2 of Theorem 16.6.2.

54.[C] Prove case 3 of Theorem 16.6.2.

55.[C] The three dimensions of a box are x, y, and z. The girth plus length are at most 165 inches. If you are free to choose which dimension is the length, which would you choose if you wanted to maximize the volume of the box? Assume $x < y < z$.

56.[C] A surface is called **closed** when it is the boundary of a region R, as a balloon surrounds the air within it. A surface is called **smooth** when it has a continuous outward unit normal vector at each point of the surface. Let S be a smooth closed surface. Show that for any point P_0 in R, there are at least two points on S such that $\overrightarrow{P_0P}$ is normal to S. NOTE: It is conjectured that if P_0 is the centroid of R, then there are at least four points on S such that P_0P is normal to S.

57.[C] Find the point P on the plane $Ax+By+Cz+D=0$ nearest the point $P_0 = (x_0, y_0, z_0)$, which is not on that plane.

(a) Find P by calculus.

(b) Find P by using the algebra of vectors. (Why is $\overrightarrow{P_0P}$ perpendicular to the plane?)

58.[C] This exercise outlines the proof of Theorem 16.6.2 in the case $f_{xx}(a,b) > 0$ and $f_{xx}(a,b)f_{yy}(a,b) - fxy^2(a,b) > 0$. Assuming that f_{xx}, f_{yy}, and f_{xy} are continuous, we know by the permanence principle that f_{xx} and $f_{xx}f_{yy} - f_{x^2y}$ remain positive throughout some disk R whose center is (a,b). The following steps show that f has a minimum (a,b) on each line L through (a,b). Let $u = \cos(\theta) + \sin(\theta)$ be a unit vector. Show that $D_u(D_uf)$ is positive throughout the part of L that lies in the dark.

(a) Show that $D_uf(a,b) = 0$.

(b) Show that $D_u(D_uf) = f_{xx}\cos^2(\theta) + 2f_{xy}\sin(\theta)\cos(\theta) + f_{yy}\sin^2(\theta)$.

(c) Show that $f_{xx}D_u(D_uf) = (f_{xx}cos(\theta) + fxy\sin(\theta))^2 + (f_{xx}f_{yy} - f_{xy}^2)\sin^2(\theta)$.

(d) Deduce from (b) that f is concave up as the part of each line through (a,b) inside the disk R.

(e) Deduce that f has a relative minimum at (a,b).

59.[C] Let $f(x)$ have period 2π and let

$$S(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} a_k\cos(kx) + \sum_{k=1}^{\infty} b_k\sin(kx)$$

be the series that minimizes the integral

$$\int_{-\pi}^{\pi} (f(x) - S(x))^2 \, dx. \qquad (16.6.10)$$

Show that $S(x)$ is the Fourier series associated with $f(x)$. NOTE: You may assume that in this case you may "differentiate past the integral sign," that is

$$\frac{\partial}{\partial y}\int_a^b g(x,y)\,dx = \int_a^b \frac{\partial g}{\partial y}\,dx.$$

The quantity in (16.6.10) measures the total squared error between $S(x)$ and $f(x)$ over the interval $[-\pi,\pi]$.

60.[C] Prove Case 2 of Theorem 16.6.2.

61.[C] Prove Case 3 of Theorem 16.6.2.

16.7 Lagrange Multipliers

Another method of finding maxima or minima of a function is due to Joseph Louis LaGrange (1736–1813). It makes use of the fact that a gradient of a function is perpendicular to the level curves (or level surfaces) of that function.

See http://en.wikipedia.org/wiki/Joseph_Louis_Lagrange.

The Essence of the Method

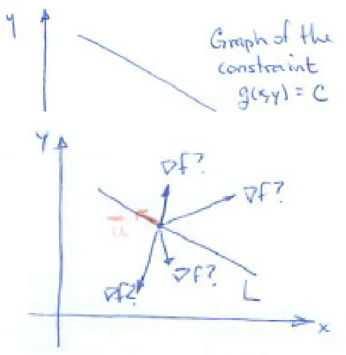


Figure 16.7.2:

We introduce the technique by considering the simplest case. Imagine that you want to find a maxima or a minima of $f(x,y)$ for points (x,y) on the line L that has the equation $g(x,y) = C$. See Figure 16.7.1.

Imagine that $f(x,y)$, for points on L has a maximum or minimum at the point (a,b). Let ∇f be the gradient of f evaluated at (a,b). What can we say about the direction of ∇f? (See Figure 16.7.2

Assume that ∇f is not perpendicular to L. Let $\mathbf{u}$ be a unit vector parallel to L. Then $D_{\mathbf{u}} f = (\nabla f)\cdot\mathbf{u}$ is not 0. If $D_{\mathbf{u}} f$ is positive then $f(x,y)$ is *increasing* in the direction $\mathbf{u}$, which is along L. In the direction $-\mathbf{u}$, $f(x,y)$ is *decreasing*. Therefore the point (a,b) could not provide either a maximum or a minimum of $f(x,y)$ for point (x,y) on L. That means ∇f must be perpendicular to L. But ∇g is perpendicular to L, since $g(x,y) = C$ is a level curve of g. Since ∇f and ∇g are parallel there must be a scalar λ such that

$$\nabla f = \lambda \nabla g \tag{16.7.1}$$

λ, *lambda*, Greek letter L.

The scalar λ is called a **Lagrange multiplier**.

EXAMPLE 1 Find the minimum of x^2y^2 on the line $x + y = 2$.

SOLUTION Since $x^2 + 2y^2$ increases without bound in both directions along the line it must have a minimum somewhere.

Here $f(x,y) = x^2 + 2y^2$ and $g(x,y) = x + y$ so

$$\nabla f = 2x\mathbf{i} + 4y\mathbf{j} \qquad \text{and} \qquad \nabla g = \mathbf{i} + \mathbf{j}$$

At the minimum, the gradients of f and g must be parallel. That is, there is a scalar λ such that

$$\nabla f = \lambda \nabla g,$$

This means

$$2x\mathbf{i} + 4y\mathbf{j} = \lambda(\mathbf{i} + \mathbf{j}). \tag{16.7.2}$$

This single vector equation leads to the 2 equations

$$\begin{cases} 2x = \lambda & \text{equating } \mathbf{i} \text{ components} \\ 4y = \lambda & \text{equating } \mathbf{j} \text{ components} \end{cases} \tag{16.7.3}$$

But we also have the constraint,

$$x + y = 2 \qquad (16.7.4)$$

From (16.7.3), $2x = 4y$ or $x = 2y$. Substituting this into (16.7.4) gives $2y+y = 2$ or $y = 2/3$, hence $x = 2y = 4/3$. The minimum is $f\left(\frac{4}{3}, \frac{2}{3}\right) = \left(\frac{4}{3}\right)^2 + \left(\frac{2}{3}\right)^2 = \frac{20}{9}$. There is no need to find λ its there just to help us compute. Its task, done, it gracefully departs. ◇

The General Method

Let us see why Lagrange's method works when the constraint not a line, but a curve. Consider this problem:

Maximize or minimize $u = f(x, y)$, given the constraint $g(x, y) = k$.

The graph of $g(x, y) = k$ is in general a curve C, as shown in Figure 16.7.3. Assume that f, *considered only on points of* C, takes a maximum (or minimum) value at the point P_0. Let C be parameterized by the vector function $\mathbf{G}(t) = x(t)\mathbf{i} + y(t)\mathbf{j}$. Let $\mathbf{G}(t_0) = \overrightarrow{OP_0}$. Then u is a function of t:

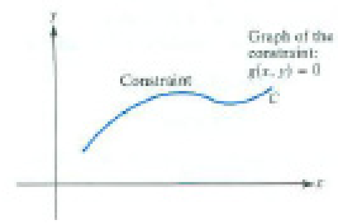


Figure 16.7.3:

$$u = f(x(t), y(t)),$$

and, as shown in the proof of Theorem 16.5 of Section 16.5,

$$\frac{du}{dt} = \nabla f \cdot \mathbf{G}'(t_0). \qquad (16.7.5)$$

Since f, considered only on C, has a maximum at $\mathbf{G}(t_0)$,

$$\frac{du}{dt} = 0 \qquad \text{at } t = 0.$$

Thus, by (16.7.5),

$$\nabla f \cdot \mathbf{G}'(t_0) = 0.$$

This means that ∇f is perpendicular to $\mathbf{G}'(t_0)$ at P_0. But ∇g, evaluated at P_0, is also perpendicular to $\mathbf{G}'(t_0)$, since the gradient ∇g is perpendicular to the level curve $g(x, y) = 0$. (We assume that ∇g is not $\mathbf{0}$.) (See Figure 16.7.4.) Thus

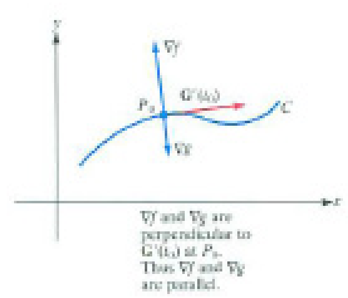


Figure 16.7.4:

∇f is parallel to ∇g.

In other words, there is a scalar λ such that $\nabla f = \lambda \nabla g$.

EXAMPLE 2 Maximize the function x^2y for points (x, y) on the unit circle $x^2 + y^2 = 1$.

SOLUTION We wish to maximize $f(x, y) = x^2y$ for points on the circle $g(x, y) = x^2 + y^2 = 1$. Then

$$\nabla f = \nabla(x^2y) = 2xy\mathbf{i} + x^2\mathbf{j}$$

and

$$\nabla g = \nabla(x^2 + y^2) = 2x\mathbf{i} + 2y\mathbf{j}$$

At an extreme point of f, $\nabla f = \lambda \nabla g$ for some scalar λ. This gives us two scalar equations:

$$2xy = \lambda(2x) \qquad \mathbf{i}\text{ component} \qquad (16.7.6)$$

$$x^2 = \lambda(2y) \qquad \mathbf{j}\text{ component} \qquad (16.7.7)$$

The third equation is the constraint,

$$x^2 + y^2 = 1. \qquad (16.7.8)$$

Since the maximum does not occur when $x = 0$, we may assume x is not 0. Dividing both sides of (16.7.6) by x, we get $2y = 2\lambda$ or $y = \lambda$. Thus (16.7.7) becomes

$$x^2 = 2y^2. \qquad (16.7.9)$$

Combining this with (16.7.8), we have

$$2y^2 + y^2 = 1$$

or

$$y^2 = \frac{1}{3}.$$

Thus

$$y = \frac{\sqrt{3}}{3} \qquad \text{or} \qquad y = -\frac{\sqrt{3}}{3}.$$

By (16.7.9),

$$x = \sqrt{2}y \qquad \text{or} \qquad x = -\sqrt{2}y.$$

There are only four points to be considered on the circle:

$$\left(\frac{\sqrt{6}}{3}, \frac{\sqrt{3}}{3}\right), \left(\frac{-\sqrt{6}}{3}, \frac{\sqrt{3}}{3}\right), \left(\frac{-\sqrt{6}}{3}, \frac{-\sqrt{3}}{3}\right), \left(\frac{\sqrt{6}}{3}, \frac{-\sqrt{3}}{3}\right).$$

At the first and second points x^2y is positive, while at the third and fourth x^2y is negative. The first two points provide the maximum value of x^2y on the circle $x^2 + y^2 = 1$, namely

$$\left(\frac{\sqrt{6}}{3}\right)^2 \frac{\sqrt{3}}{3} = \frac{2\sqrt{3}}{9}.$$

The third and fourth points provide the minimum value of x^2y namely,

$$\frac{-2\sqrt{3}}{9}.$$

◇

More Variables

In the preceding examples we examined the maximum and minimum of $f(x, y)$ on a curve $g(x, y) = k$. But the same method works for dealing with extreme values of $f(x, y, z)$ on a surface $g(x, y, z) = k$. If f has, say, a minimum at (a, b, c), then it does on any level curve on the surface $g(x, y, z) = k$. Thus ∇f is perpendicular to any curve on the surface through P. But so is ∇g. Thus ∇f and ∇g are parallel, and there is a scalar λ such that the $\nabla f = \lambda \nabla g$. So we will have four scalar equations: three from the vector equation $\nabla f = \lambda \nabla g$ and one from the constraint $g(x, y, z) = k$. That gives four equations in four unknowns, x, y, z and λ, but it is not necessary to find λ though it may be useful to determine it. Solving these four simultaneous equations may not be feasible. However, the exercises in this section lead to fairly simple equations that are relatively easy to solve.

EXAMPLE 3 Find the rectangle box with the largest volume, given that its surface area is 96 square feet.

SOLUTION Let the three dimensions be x, y and z and the volume be V, which equals xyz. The surface area is $2xy + 2xz + 2yz$. See Figure 16.7.5.

We wish to maximize $V(x, y, z) = xyz$ subject to the constraint

$$g(x, y, z) = 2xy + 2xz + 2yz = 96. \qquad (16.7.10)$$

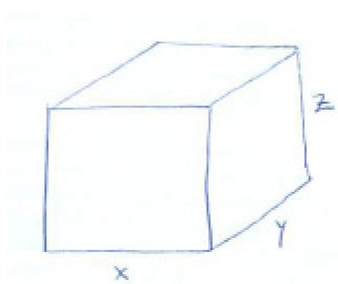


Figure 16.7.5:

Now

$$\nabla V = yz\mathbf{i} + xz\mathbf{j} + xy\mathbf{k}$$

and

$$\nabla g = (2y + 2z)\mathbf{i} + (2x + 2z)\mathbf{j} + (2x + 2y)\mathbf{k}.$$

The vector equation $\nabla V = \lambda \nabla g$ provides three scalar equations

$$\begin{aligned} yz &= \lambda(2y+2z) \\ xz &= \lambda(2x+2z) \\ xy &= \lambda(2x+2y) \end{aligned}$$

The fourth equation is the constraint,

$$2xy + 2xz + 2yz = 96.$$

Solving for λ in (16.7.11) and in (16.7.11), and equating the results gives

$$\frac{yz}{2y+2z} = \frac{xz}{2x+2z}.$$

Why not?

Since z will not be 0, we have

$$\frac{y}{2y+2z} = \frac{x}{2x+2z}.$$

Clearing denominators gives

$$\begin{aligned} 2xy + 2yz &= 2xy + 2xz \\ 2yz &= 2xz. \end{aligned}$$

Since $z \neq 0$, we reach the conclusion that

$$x = y.$$

Since x, y and z play the same roles in both the volume xyz and in the surface area, $2(xy + xz + yz)$, we conclude also that

$$x = z.$$

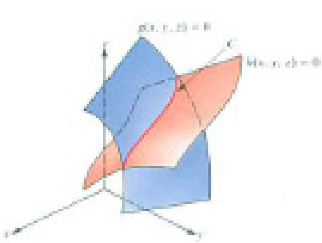
Then $x = y = z$. The box of maximum volume is a cube.

To find its dimensions we return to the constraint, which tells us that $6x^2 = 96$ or $x = 4$. Hence y and z are 4 also. ◇

More Constraints

Lagrange multipliers can also be used to maximize $f(x, y, z)$ subject to more than one constraint; for instance, the constraints may be

$$g(x,y,z) = k_1 \qquad \text{and} \qquad h(x,y,z) = k_2. \qquad (16.7.11)$$



The two surfaces (16.7.11) in general meet in a curve C, as shown in Figure 16.7.6. Assume that C is parameterized by the function $\mathbf{G}$. Then at a maximum (or minimum) of f at a point $P_0(x_0, y_0, z_0)$ on C,

$$\nabla f \cdot \mathbf{G}'(t_0) = 0.$$

Thus ∇f, evaluated at P_0, is perpendicular to $\mathbf{G}'(t_0)$. But ∇g and ∇h, being normal vectors at P_0 to the level surfaces $g(x, y, z) = K_1$ and $h(x, y, z) = K_2$, respectively, are both perpendicular to $\mathbf{G}'(t_0)$. Thus

∇f, ∇g, and ∇h are all perpendicular to $\mathbf{G}'(t_0)$ at (x_0, y_0, z_0).

(See Figure 16.7.7.) Consequently, ∇f lies in the plane determined by the vectors ∇g and ∇h (which we assume are not parallel). Hence there are scalars λ and μ such that

$$\nabla f = \lambda \nabla g + \mu \nabla h.$$

This vector equation provides three scalar equations in λ, μ, x, y, z. The two constraints give two more equations. All told: five equations in five unknowns. (Of course we find λ and μ only if they assist the algebra.)

mu, *mew*, is Greek for the letter M.

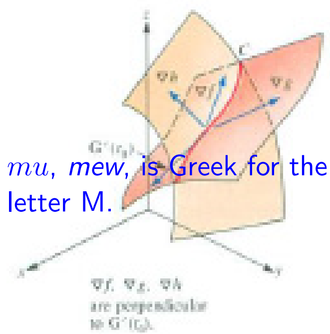


Figure 16.7.7:

A rigorous development of the material in this section belongs in an advanced calculus course. If a maximum occurs at an endpoint of the curves in question or if the two surfaces do not meet in a curve or if the ∇g and ∇h are parallel, this method does not apply. We will content ourselves by illustrating the method with an example in which there are two constraints.

EXAMPLE 4 Minimize the quantity $x^2 + y^2 + z^2$ subject to the constraints $x + 2y + 3z = 6$ and $x + 3y + 9z = 9$.

SOLUTION There are three variables and two constraints. Each of the two constraints mentioned describes a plane. Thus the two constraints together describe a *line*. The function $x^2 + y^2 + z^2$ is the square of the distance from (x, y, z) to the origin. So the problem can be rephrased as "How far is the origin from a certain line?" (It could be solved by vector algebra. See Exercises 19 and 20.) When viewed this way, the problem certainly has a solution; that is, there is clearly a minimum.

In this case

$$f(x, y, z) = x^2 + y^2 + z^2 \qquad (16.7.12)$$
$$g(x, y, z) = x + 2y + 3z \qquad (16.7.13)$$
$$h(x, y, z) = x + 3y + 9z. \qquad (16.7.14)$$

Thus

$$\nabla f = 2x\mathbf{i} + 2y\mathbf{j} + 2z\mathbf{k} \qquad (16.7.15)$$
$$\nabla g = \mathbf{i} + 2\mathbf{j} + 3\mathbf{k} \qquad (16.7.16)$$
$$\nabla h = \mathbf{i} + 3\mathbf{j} + 9\mathbf{k}. \qquad (16.7.17)$$

There are constants λ and μ so

$$\nabla f = \lambda \nabla g + \mu \nabla h.$$

Therefore, the five equations for x, y, z, λ, and μ are

$$2x = \lambda + \mu \quad (16.7.18)$$
$$2y = 2\lambda + 3\mu \quad (16.7.19)$$
$$2z = 3\lambda + 9\mu \quad (16.7.20)$$
$$x + 2y + 3z = 6 \quad (16.7.21)$$
$$x + 3y + 9z = 9 \quad (16.7.22)$$

One way is to use software programs that solve simultaneous linear equations.

There are several ways to solve these equations.

One way is to use the first three of the five equations: to express x, y, and z in terms of λ and μ. Then substitute these values in the last two equations, getting an old friend from high school "two simultaneous equations in two unknowns"

By (16.7.18), (16.7.19), and (16.7.20),

$$x = \frac{\lambda + \mu}{2}, \qquad y = \frac{2\lambda + 3\mu}{2}, \qquad z = \frac{3\lambda + 9\mu}{2}.$$

Equations (16.7.21) and (16.7.22) then become

$$\frac{\lambda + \mu}{2} + \frac{2(2\lambda + 3\mu)}{2} + \frac{3(3\lambda + 9\mu)}{2} = 6$$

and

$$\frac{\lambda + \mu}{2} + \frac{3(2\lambda + 3\mu)}{2} + \frac{9(3\lambda + 9\mu)}{2} = 9,$$

which simplify to

$$14\lambda + 34\mu = 12 \quad (16.7.23)$$

and
$$34\lambda + 91\mu = 18. \quad (16.7.24)$$

Solving (16.7.23) and (16.7.24) gives

$$\lambda = \frac{240}{59} \qquad \mu = -\frac{78}{59}.$$

Thus

$$x = \frac{\lambda + \mu}{2} = \frac{81}{59} \approx 1.37288,$$
$$y = \frac{2\lambda + 3\mu}{2} = \frac{123}{59} \approx 2.08475,$$
$$z = \frac{3\lambda + 9\mu}{2} = \frac{9}{59} \approx 0.15254.$$

The minimum of $x^2 + y^2 + z^2$ is this

$$\left(\frac{81}{59}\right)^2 + \left(\frac{123}{59}\right)^2 + \left(\frac{9}{59}\right)^2 = \frac{21,771}{3,481} = \frac{369}{59} \approx 6.24542.$$

Since there is no maximum, this must be a minimum. Why?

◇

In Example 4 there were three variables, x, y, and z, and two constraints. There may, in some cases, be many variables, x_1, x_2, $\ldots x_n$, and many constraints. If there are m constraints, g_1, $g_2 \ldots g_m$ introduce Lagrange multipliers λ_1, λ_2, $\ldots \lambda_m$, one for each constraints. So there would be $m + n$ equations, n from the equation

$$\nabla f = \lambda_1 \nabla g_1 + \lambda_2 \nabla g_2 + \cdots + \lambda_m \nabla g_m$$

and m more equations from the m constraints. There would be $m + n$ unknowns, λ_1, λ_2, $\ldots$, λ_m, x_1, x_2, $\ldots$, x_n.

Summary

The basic idea of Lagrange multipliers is that if $f(x, y, z)$ (or $f(x, y)$) has an extreme value on a curve that lies on the surface $g(x, y, z) = C$ (or the curve $g(x, y) = k$), then ∇f and ∇g are both perpendicular to the curve at the point where the extreme value occurs. If there is only one constraint, then ∇f and ∇g are parallel. If there are two constraints $g(x, y, z) = k_1$ and $h(x, y, z) = k_2$, then ∇f lies on the plane of ∇g and ∇h. In the first case there is a scalar λ such that $\nabla f = \lambda \nabla g$. In the second case, there are scalars λ and μ such that $\nabla f = \lambda \nabla g + \mu \nabla h$. These vector equations, together with the constraints, provide simultaneous scalars equations, which must then be solved.

EXERCISES for Section 16.7 *Key:* R–routine, M–moderate, C–challenging

In the exercises use Lagrange multipliers unless otherwise suggested.

1.[R] Maximize xy for points on the circle $x^2+y^2=4$.

2.[R] Minimize x^2+y^2 for points on the line $2x+3y=6$.

3.[R] Minimize $2x+3y$ on the portion of the hyperbola $xy=1$ in the first quadrant.

4.[R] Maximize $x+2y$ on the ellipse $x^2+y^2=8$.

5.[R] Find the largest area of all rectangles whose perimeters are 12 centimeters.

6.[R] A rectangular box is to have a volume of 1 cubic meter. Find its dimensions if its surface area is minimal.

7.[R] Find the point on the plane $x+2y+3z=6$ that is closest to the origin. HINT: Minimize the square of the distance in order to avoid square roots.

8.[R] Maximize $x+y+2z$ on the sphere $x^2+y^2z^2=9$.

9.[R] Minimize the distance from (x,y,z) to $(1,3,2)$ for points on the plane $2x+y+z=5$.

10.[R] Find the dimensions of the box of largest volume whose surface area is to be 6 square inches.

11.[R] Maximize $x^2y^2z^2$ subject to the constraint $x^2+y^2+z^2=1$.

12.[R] Find the points on the surface $xyz=1$ closest to the origin.

13.[R] Minimize $x^2+y^2+z^2$ on the line common to the two planes $x+2y+3z=0$ and $2x+3y+z=4$.

14.[R] The plane $2y+4z-5=0$ meets the cone $z^2=4(x^2+y^2)$ in a curve. Find the point on this curve nearest the origin.

In Exercises 15 to 18 solve the given exercise in Section 16.5 by Lagrange multipliers.

15.[R] Exercise 25

16.[R] Exercise 26

17.[R] Exercise 29

18.[R] Exercise 30

19.[R] Solve Example 4 by vector algebra.

20.[R] Solve Exercise 13 by vector algebra.

21.[R]

(a) Sketch the elliptical paraboloid $z=x^2+2y^2$.

(b) Sketch the plane $x+y+z=1$.

(c) Sketch the intersection of the surfaces in (a) and (b).

(d) Find the highest point on the intersection in (c).

22.[R]

(a) Sketch the ellipsoid $x^2+y^2/4+z^2/9=1$ and the point $P(2,1,3)$.

(b) Find the point Q on the ellipsoid that is nearest P.

(c) What is the angle between PQ and the tangent plane at Q?

23.[R]

(a) Sketch the hyperboloid $x^2-y^2/4-z^4/9=1$. (How many sheets does it have?)

(b) Sketch the point $(1, 1, 1)$. (Is it "inside" or "outside" the hyperboloid?)

(c) Find the point on the hyperboloid nearest P.

24.[R] Maximize $x^3 + y^3 + 2z^3$ on the intersection of the surfaces $x^2+y^2+z^2 = 4$ and $(x-3)^2+y^2+z^2 = 4$.

25.[R] Show that a triangle in which the product of the sines of the three angles is maximized is equilateral. HINT: Use Lagrange multipliers.

26.[R] Solve Exercise 25 by labeling the angles x,y, and $\pi - x - y$ and minimizing a function of x and y by the method of Section 16.6.

27.[R] Maximize $x+2y+3z$ subject to the constraints $x^2 + y^2 + z^2 = 2$ and $x + y + z = 0$.

28.[C]

(a) Maximize $x_1x_2, \cdots x_n$ subject to the constraints that $\sum_{i=1}^{n} x_i = 1$ and all $x_i \geq 0$.

(b) Deduce that for nonnegative numbers a_1, a_2, ..., a_n, $\sqrt[n]{a_1a_2\cdots a_n} \leq (a_1 + a_2 + \cdots + a_n)/n$. (The **geometric mean** is less than or equal to the **arithmetic mean**.)

29.[C]

(a) Maximize $\sum_{i=1}^{n} x_iy_i$ subject to the constraints $\sum_{i=1}^{n} x_i^2 = 1$ and $\sum_{i=1}^{n} y_i^2 = 1$.

(b) Deduce that for any numbers a_1, a_2, ..., a_n and $b_1, b_2, \ldots b_n$, $\sum_{i=1}^{n} a_ib_i \leq \left(\sum_{i=1}^{n} a_i^2\right)^{1/2}\left(\sum_{i=1}^{n} b_i^2\right)^{1/2}$, which is called the Schwarz inequality. HINT: Let $x_i = \frac{a_i}{(\sum_{i=1}^{n} a_i^2)^{1/2}}$ and $y_1 = \frac{b_i}{(\sum_{i=1}^{n} b_i^2)^{1/2}}$.

(c) How would you justify the inequality in (b), for $n = 3$, by vectors?

30.[C] Let $a_1, a_2 \ldots a_n$ be fixed nonzero numbers. Maximize $\sum_{i=1}^{n} a_ix_i$ subject to $\sum_{i=1}^{n} x_i^2 = 1$.

31.[C] Let p and q be positive numbers that satisfy the equation $1/p+1/q = 1$. Obtain Holder's inequality for nonnegative numbers a_i and b_i,

$$\sum_{i=1}^{n} a_ib_i \leq \left(\sum_{i=1}^{n} a_i^p\right)^{1/p}\left(\sum_{i=1}^{n} b_i^q\right)^{1/q},$$

as follows.

(a) Maximize $\sum_{i=1}^{n} x_iy_i$ subject to $\sum_{i=1}^{n} x_i^p = 1$ and $\sum_{i=1}^{n} y_i^q = 1$.

(b) By letting $x_i = \frac{a_i}{(\sum_{i=1}^{n} a_i^p)^{1/p}}$ and $y_i = \frac{b_i}{(\sum_{i=1}^{n} b_i^q)^{1/q}}$, obtain Holder's inequality.

Note that Holder's inequality, with $p = 2$ and $q = 2$, reduces to the Schwarz inequality in Exercise 29.

32.[C] A consumer has a budget of B dollars and my purchase n different items. The price of the ith item is p_1 dollars. When the consumer buys x_1 units of the ith item, the total cost is $\sum_{i=1}^{n} p_ix_i$. Assume that $\sum_{i=1}^{n} p_ix_i = B$ and that the consumer wishes to maximize her utility $u(x_1, x_2 \ldots x_n)$.

(a) Show that when $x_1, \ldots, x_n$, are chosen to maximize utility, then

$$\frac{\partial u/\partial x_i}{p_i} = \frac{\partial u/\partial x_j}{p_j}.$$

(b) Explain the result in (a) using just economic intuition. HINT: Consider a slight change in x_i and x_j, with the other x_k's held fixed.

33.[C] The following is quoted from Colin W. Clark in *Mathematical Bioeconomics*, Wiley, New York, 1976:

> [S]uppose there are N fishing grounds. Let $H^i = H^i(R^i, E^i)$ denotes the production function for the total harvest H^i on the ith ground as a function of the recruited stock level R^i and effort E^i on the ith ground. The problem is to determine the least total cost $\sum_{i=1}^{N} c_i E^i$ at which a given total harvest $H = \sum_{i=1}^{n} H^i$ can be achieved. This problem can be easily solved by Lagrange multipliers. The result is simply
>
> $$\frac{1}{c_i}\frac{\partial H^i}{\partial E^i} = \text{constant}$$
>
> [independent of i].

Verify his assertion. The c_i's are constants. The superscripts name the functions; they are not exponents.

34.[C] (*Computer science*) This exercise is based on J. D. Ullman, *Principles of Database Systems*, pp. 82–83, Computer Science Press, Potomac, Md., 1980. It arises in the design of efficient "bucket" s is a particular way of rearrangin database.) Let p_1, p_2, $\dots$, p_k and stants. Let b_1, b_2, $\dots$, b_k be k no satisfying $\sum_{j=1}^{k} b_j = B$. The quan represents the expected search tin $b_1, b_2, \dots, b_k$ does the method of L suggest provide the minimum expe

35.[C] Assume that $f(x, y, z)$ ha at P_0 on the level surface $g(x, y, z)$

(a) Why is ∇g evaluated at P_0 p surface at P_0?

(b) Why is ∇f evaluated at P_0 p surface at P_0?

36.[C] Solve Example 35 by vec algebra).

16.8 What Everyone Who Will Study Thermodynamics Needs to Know

The basic equations of thermodynamics follow from the Chain Rule and the equality of the mixed partial derivatives. We will describe the mathematics within the thermodynamics context.

Review the Chain Rule, if necessary.

Implications of The Chain Rule

We start with a function of three variables, $f(x, y, z)$, which we assume has first partial derivatives

$$\left.\frac{\partial f}{\partial x}\right|_{y,z} \qquad \left.\frac{\partial f}{\partial y}\right|_{x,z} \qquad \left.\frac{\partial f}{\partial z}\right|_{x,y}.$$

The subscripts denote the variables held fixed.

This notation is standard practice in thermodynamics, though it offends some mathematicians.

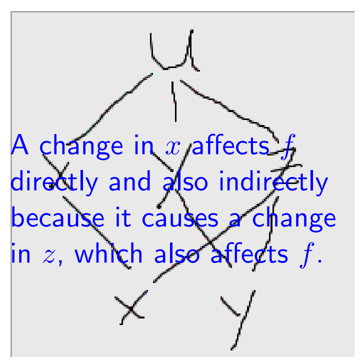
Without this explicit reminder it is necessary to remember the other variables. At this point this is not difficult. But, when additional information is included, it can become more difficult to keep track of all of the variables in the problem.

Now assume that z is a function of x and y, $z = g(x, y)$. Then $f(x, y, z) = f(x, y, g(x, y))$ is a function of only two variables. This new function we name $h(x, y)$: $h(x, y) = f(x, y, g(x, y))$. There are only two first partial derivatives of h:

$$\left.\frac{\partial h}{\partial x}\right|_{y} \quad \text{and} \quad \left.\frac{\partial h}{\partial y}\right|_{x}.$$

Let the value of $f(x, y, z)$ be called u, $u = f(x, y, z)$. But x, y, and z are functions of x and y: $x = x$, $y = y$, and $z = g(x, y)$.

Figure 16.8.1 provides a pictorial view of the relationship between the different variables. Both x and y appear as middle and independent variables. We have $u = f(x, y, z)$ and also $u = h(x, y)$. By the Chain Rule Then



A change in x affects f directly and also indirectly because it causes a change in z, which also affects f.

Figure 16.8.1:

$$\left.\frac{\partial h}{\partial x}\right|_{y} = \left.\frac{\partial f}{\partial x}\right|_{y,z}\left.\frac{\partial x}{\partial x}\right|_{y} + \left.\frac{\partial f}{\partial y}\right|_{x,z}\left.\frac{\partial y}{\partial x}\right|_{y} + \left.\frac{\partial f}{\partial z}\right|_{x,y}\left.\frac{\partial g}{\partial x}\right|_{y}.$$

Since x and y are independent variables, $\partial x/\partial x = 1$ and $\partial y/\partial x = 0$ and we have

$$\left.\frac{\partial h}{\partial x}\right|_{y} = \left.\frac{\partial f}{\partial x}\right|_{y,z} + \left.\frac{\partial f}{\partial z}\right|_{x,y}\left.\frac{\partial g}{\partial x}\right|_{y}, \tag{16.8.1}$$

or simply

$$\frac{\partial h}{\partial x} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial z}\frac{\partial g}{\partial x}. \tag{16.8.2}$$

When the subscripts are omitted we have to look back at the definitions of f, g, and h to see which variables are held fixed.

EXAMPLE 1 Let's check (16.8.2) when

$$f(x,y,z) = x^2y^3z^5 \quad \text{and} \quad g(x,y) = 2x + 3y.$$

SOLUTION We have $h(x,y) = f(x,y,g(x,y)) = x^2y^3(2x+3y)^5$. Then $\frac{\partial f}{\partial x} = 2xy^3z^5$ and $\frac{\partial f}{\partial z} = 5x^2y^3z^4$. Also $\frac{\partial g}{\partial x} = 2$.

Computing $\partial h/\partial x$ directly gives

$$\begin{aligned} \frac{\partial h}{\partial x} &= \frac{\partial}{\partial x}\left(x^2y^3(2x+3y)^5\right) \\ &= y^3\frac{\partial}{\partial x}\left(x^2(2x+3y)^5\right) \\ &= y^3\left(2x(2x+3y)^5 + x^2\left(5(2x+3y)^4(2)\right)\right) \\ &= 2xy^3(2x+3y)^5 + 10x^2y^3(2x+3y)^4. \end{aligned} \tag{16.8.3}$$

On the other hand, by (16.8.2), we have

$$\begin{aligned} \frac{\partial h}{\partial x} &= \frac{\partial f}{\partial x} + \frac{\partial f}{\partial z}\frac{\partial g}{\partial x} \\ &= 2xy^3z^5 + (5x^2y^3z^4)(2) \\ &= 2xy^3(2x+3y)^5 + 10x^2y^3(2x+3y)^4, \end{aligned}$$

which agrees with (16.8.3). ◇

What If $z = g(x,y)$ Makes $f(x,y,z)$ Constant?

Next, assume that when z is replaced by $g(x,y)$, the function $h(x,y) = f(x,y,g(x,y))$ is constant: $h(x,y) = f(x,y,g(x,y)) = C$. This happens when we use the equation $f(x,y,z) = C$ to determine z implicitly as a function of x and y.

Then

$$\left.\frac{\partial h}{\partial x}\right|_y = 0 \quad \text{and} \quad \left.\frac{\partial h}{\partial y}\right|_x = 0.$$

(16.8.4) will be the foundation for deriving (16.8.9) and (16.8.10), key mathematical relationships used in thermodynamics.

In this case, which occurs frequently in thermodynamics, (16.8.1) becomes

$$0 = \left.\frac{\partial f}{\partial x}\right|_{y,z} + \left.\frac{\partial f}{\partial z}\right|_{x,y}\left.\frac{\partial g}{\partial x}\right|_y. \tag{16.8.4}$$

Solving (16.8.4) for $\left.\frac{\partial g}{\partial x}\right|_y$ we obtain

$$\left.\frac{\partial g}{\partial x}\right|_y = \frac{-\left.\frac{\partial f}{\partial x}\right|_{y,z}}{\left.\frac{\partial f}{\partial z}\right|_{x,y}}. \tag{16.8.5}$$

Equation (16.8.5) expresses the partial derivative of $g(x, y)$ with respect to x in terms of the partial derivatives of the original function $f(x, y, z)$.

EXAMPLE 2 Let $f(x, y, z) = x^3y^5z^7$. Define $g(x, y)$ implicitly by $x^3y^5(g(x, y))^7 = 1$. That is, $g(x, y) = x^{-3/7}y^{-5/7}$. Verify (16.8.5).

SOLUTION First of all, $\left.\frac{\partial g}{\partial x}\right|_y = \frac{-3}{7}x^{-10/7}y^{-5/7}$. Then

$$\left.\frac{\partial f}{\partial x}\right|_{y,z} = 3x^2y^5z^7 \quad \text{and} \quad \left.\frac{\partial f}{\partial z}\right|_{x,x} = 7x^3y^5z^6.$$

Substituting in (16.8.5), we have

$$\begin{aligned}
\frac{-\left.\frac{\partial f}{\partial x}\right|_{y,z}}{\left.\frac{\partial f}{\partial z}\right|_{x,y}} &= \frac{-\left(3x^2y^5z^7\right)}{7x^3y^5z^6} \\
&= -\frac{3}{7}x^{-1}z \\
&= -\frac{3}{7}x^{-1}x^{-3/7}y^{-5/7} && \text{because } x^3y^5z^7 = 1 \\
&= -\frac{3}{7}x^{-10/7}y^{-5/7} \\
&= \left.\frac{\partial g}{\partial x}\right|_y && \text{so (16.8.5) is satisfied.}
\end{aligned}$$

◇

The Reciprocity Relations

In a thermodynamics text you will see equations of the form

$$\left.\frac{\partial x}{\partial z}\right|_y = \frac{1}{\left.\frac{\partial z}{\partial x}\right|_y}. \tag{16.8.6}$$

We will explain where this equation comes from, presenting the mathematical details often glossed over in the applied setting. There is a function $f(x, y, z)$ with constant value C, $f(x, y, z) = C$. It is assumed that this equation determines z as a function of x and y, or, similarly, determines x as a function of y and z, or y as a function of x and z. There are six first partial derivatives:

See Exercise 5.

$$\left.\frac{\partial z}{\partial x}\right|_y, \quad \left.\frac{\partial z}{\partial y}\right|_x, \quad \left.\frac{\partial x}{\partial y}\right|_z, \quad \left.\frac{\partial x}{\partial z}\right|_y, \quad \left.\frac{\partial y}{\partial x}\right|_z, \quad \left.\frac{\partial y}{\partial z}\right|_x. \tag{16.8.7}$$

An equation analogous to (16.8.5) holds for each of them. For instance,

$$\left.\frac{\partial x}{\partial z}\right|_y = \frac{-\left.\frac{\partial f}{\partial z}\right|_{x,y}}{\left.\frac{\partial f}{\partial x}\right|_{y,z}}. \tag{16.8.8}$$

Combining (16.8.5) and (16.8.8) verifies that

This is to be expected, for $\frac{\Delta z}{\Delta x}$ is the reciprocal of $\frac{\Delta x}{\Delta z}$.

$$\left.\frac{\partial x}{\partial z}\right|_y = \frac{1}{\left.\frac{\partial z}{\partial x}\right|_y}. \tag{16.8.9}$$

Equation (16.8.9) is an example of a **reciprocity relation**: The partial derivative of one variable with respect to a second variable is the reciprocal of the partial derivative of the second variable with respect to the first variable.

EXAMPLE 3 Let $f(x, y, z) = 2x + 3y + 5z = 12$. Verify that $\partial z/\partial x$ is the reciprocal of $\partial x/\partial z$.

SOLUTION Since $2x + 3y + 5z = 12$, $z = (12 - 2x - 3y)/5$. Then $\partial z/\partial x = -2/5$.

Also, $x = (12 - 3y - 5z)/2$, so $\partial x/\partial z = -5/2$, which is, as predicted, the reciprocal of $\partial z/\partial x$. ◇

The Cyclic Relations

With the aid of equations like (16.8.8) it is easy to establish the surprising relation

The **Cyclic Relation**, also known as the Triple Product Rule, the Cyclic Chain Rule, or Euler's Chain Rule. See http://en.wikipedia.org/wiki/Triple_

$$\left.\frac{\partial x}{\partial y}\right|_z \left.\frac{\partial y}{\partial z}\right|_x \left.\frac{\partial z}{\partial x}\right|_y = -1. \tag{16.8.10}$$

Equation (16.8.10) results from the use of three versions of (16.8.8). The left-hand side of (16.8.10) can be expressed as

$$\left(\frac{-\left.\frac{\partial f}{\partial y}\right|_{x,z}}{\left.\frac{\partial f}{\partial x}\right|_{y,z}}\right)\left(\frac{-\left.\frac{\partial f}{\partial z}\right|_{x,y}}{\left.\frac{\partial f}{\partial y}\right|_{x,z}}\right)\left(\frac{-\left.\frac{\partial f}{\partial z}\right|_{x,y}}{\left.\frac{\partial f}{\partial x}\right|_{y,z}}\right) \tag{16.8.11}$$

Cancellation reduces (16.8.11) to -1.

EXAMPLE 4 Let $f(x, y, z) = 2x+3y+5z = 12$. This equation determines implicitly each of the variables in terms of the two others. Verify (16.8.10) in this case.

SOLUTION By the equation $2x + 3y + 5z = 12$,

$$x = \frac{12 - 3y - 5z}{2} \qquad y = \frac{12 - 2x - 5z}{3} \qquad z = \frac{12 - 2x - 3y}{5}$$

Then $\partial x/\partial y = -3/2$, $\partial y/\partial z = -5/3$, and $\partial z/\partial x = -2/5$, and we have

$$\frac{\partial x}{\partial y}\,\frac{\partial y}{\partial z}\,\frac{\partial z}{\partial x} = \left(\frac{-3}{2}\right)\left(\frac{-5}{3}\right)\left(\frac{-2}{5}\right) = -1$$

◇

If two of the three partial derivatives in (16.8.10) are easy to calculate, then we can use (16.8.10) to find the third, which may otherwise be hard to calculate. We illustrate this use of the cyclic relationship with an example from thermodynamics. In this context T denotes temperature, p, pressure, and v the mass per unit volume.

v is the reciprocal of density

Equations (16.8.4, (16.8.9), and (16.8.10) are the 15sential mathematical relationships used in thermodynamics. We now show their use in a few typical thermodynamics problems.

EXAMPLE 5 In van der Waal's equation p, T, and v are all related by the relation

$$p = \frac{RT}{v - b} - \frac{a}{v^2}; \tag{16.8.12}$$

van der Waal's equation is only one example of an **equation of state**. See also Exercises 11 and 12.

Exercises 13 and 14 describe other ways to solve Example 5.

R, a and b are constants. Use a cyclic relation to find $(\partial v/\partial T)_p$.

SOLUTION We use the cyclic relation

$$\left.\frac{\partial v}{\partial T}\right|_p \left.\frac{\partial T}{\partial p}\right|_v \left.\frac{\partial p}{\partial v}\right|_T = -1. \tag{16.8.13}$$

Looking at (16.8.12), we see that $(\partial p/\partial T)_v$ is easier to calculate than $(\partial T/\partial p)_v$. So (16.8.13) becomes

$$\frac{\left.\frac{\partial v}{\partial T}\right|_p \left.\frac{\partial p}{\partial v}\right|_T}{\left.\frac{\partial p}{\partial T}\right|_v} = -1$$

and therefore

$$\left.\frac{\partial v}{\partial T}\right|_p = -\frac{\left.\frac{\partial p}{\partial T}\right|_v}{\left.\frac{\partial p}{\partial v}\right|_T}. \tag{16.8.14}$$

Since p is given as a function of v and T, it is easy to calculate the numerator and denominator in (16.8.14):

$$\left(\frac{\partial p}{\partial T}\right)_v = \frac{R}{v-b} \quad \text{and} \quad \left(\frac{\partial p}{\partial v}\right)_T = \frac{-RT}{(v-b)^2} + \frac{2a}{v^3}.$$

Thus, by (16.8.14),

$$\left(\frac{\partial v}{\partial T}\right)_p = \frac{-R/(v-b)}{-RT/(v-b)^2 + 2a/v^3}.$$

◇

Using the Equality of the Mixed Partial Derivatives

Having shown how the Chain Rule provides some of the basic equations in thermodynamics, let us show how the equality of the mixed partials leads to other basic equations.

We resume our consideration of a thermodynamic process in which the pressure is denoted by P, the temperature by T, and the volume per unit mass by v. Other common variables are

u	thermal energy per unit mass
s	entropy per unit mass
a	Helmholtz free energy per unit mass
g	Gibbs free energy per unit mass
h	enthalpy per unit mass

That is a total of 8 variables of interest. If they were independent, the possible states would be part of an eight-dimensional space. However, they are very *inter*dependent. In fact *any two* determine all the others.

For instance, u may be viewed as a function of s and v, and we have $\left.\frac{\partial u}{\partial s}\right|_v$, which is the *definition* of temperature, T. Thermodynamic texts either state or derive the "Gibbs relation"

When you look at your thermometer, remember that you are gazing at the value of a partial derivative.

$$du = T\ ds - P\ dv.$$

This equation involving differentials tells us that u is viewed as a function of s and v, and that

$$\left.\frac{\partial u}{\partial s}\right|_v = T \quad \text{and} \quad \left.\frac{\partial u}{\partial v}\right|_s = -P.$$

Equating the mixed second partial derivatives then gives us

$$\begin{aligned}
\frac{\partial^2 u}{\partial v \partial s} &= \frac{\partial^2 u}{\partial s \partial v} && \text{equality of mixed partials of } u(s,v) \\
\frac{\partial}{\partial v}\left(\frac{\partial u}{\partial s}\right) &= \frac{\partial}{\partial s}\left(\frac{\partial u}{\partial v}\right) && \\
\left.\frac{\partial T}{\partial v}\right|_s &= \left.\frac{\partial (-P)}{\partial s}\right|_v && \text{because } \left.\frac{\partial u}{\partial s}\right|_v = T \text{ and } \left.\frac{\partial u}{\partial v}\right|_s = -P \\
\left.\frac{\partial T}{\partial v}\right|_s &= -\left.\frac{\partial P}{\partial s}\right|_v .
\end{aligned}$$

Several thermodynamic statements that equate two partial derivatives are obtained this way. The starting point is an equation of the form

In other contexts we will say that $dz = Mdx + Ndy$ is an **exact differential**.

$$dz = M\ dx + N\ dy$$

where M is $\left.\frac{\partial z}{\partial x}\right|_y$ and N is $\left.\frac{\partial z}{\partial y}\right|_x$. Then, because

$$\frac{\partial z}{\partial x \partial y} = \frac{\partial z}{\partial y \partial x},$$

it is found that

$$\left.\frac{\partial M}{\partial y}\right|_x = \left.\frac{\partial N}{\partial x}\right|_y .$$

Summary

We showed how the Chain Rule in the special case where an intermediate variable is also a final variable justifies certain identities, namely, the *reciprocal* and *cyclic relations* used in thermodynamics. Then we showed how the equality of the mixed partial derivatives is used to derive other equations linking various partial derivatives.

EXERCISES for Section 16.8 *Key:* R–routine, M–moderate, C–challenging

1.[R] Let $u = x^2 + y^2 + z^2$ and let $z = x + y$.

(a) The symbol $\frac{\partial u}{\partial x}$ has two interpretations. What are they?

(b) Evaluate $\frac{\partial u}{\partial x}$ in both cases identified in (a).

2.[R] Let $z = rst$ and let $r = st$.

(a) The symbol $\frac{\partial z}{\partial t}$ has two interpretations. What are they?

(b) Evaluate $\frac{\partial z}{\partial t}$ in both cases identified in (a).

3.[R] Let $u = f(x, y, z)$ and $z = g(x, y)$. Then u is indirectly a function of x and of y. Express $\left.\frac{\partial u}{\partial x}\right|_y$ in terms of partial derivatives of f. (Supply all the steps.)

4.[R] Assume that the equation $f(x, y, z) = C$, a constant, determines x as a function of y and z: $x = h(y, z)$. Express $\left.\frac{\partial x}{\partial y}\right|_z$ in terms of partial derivatives of f. (Supply all the steps.)

5.[R] What is the product of the six partial derivatives in (16.8.7)?

6.[R] Using the function f from Example 2, verify the analog of (16.8.8) for $\left.\frac{\partial z}{\partial y}\right|_x$.

7.[R] Let $f(x, y, z) = 2x + 4y + 3z$. The equation $f(x, y, z) = 7$ determines any variable as a funciton of the other two. Verify (16.8.8), where z is viewed as a function of x and y.

8.[R] Obtain the cyclic relation

$$\left.\frac{\partial x}{\partial z}\right|_y \left.\frac{\partial z}{\partial y}\right|_x \left.\frac{\partial y}{\partial x}\right|_z = -1.$$

HINT: Duplicate the steps leading to (16.8.10).

9.[R] Verify (16.8.10) in the case $f(x, y, z) = x^3y^5z^7 = 1$.

10.[R] Verify (16.8.10) in the case $f(x, y, z) = 2x + 4y + 3z = 7$.

11.[R] The equation of state for an ideal gas is $pv = RT$. Find $(\partial v/\partial T)_p$.

12.[R] The Redlich-Kwang equation

$$p = \frac{RT}{v - b} - \frac{a}{v(v + b)T^{1/2}}.$$

is an improvement upon the van der Waal's equation of state (16.8.12) for gases and liquids. Find $(\partial v/\partial T)_p$. NOTE: Do a Google search for "Redlich Kwang equation", or visit `http://en.wikipedia.org/wiki/Equation_of_state`.

13.[R] Find $(\partial v/\partial T)_p$ in Example 5 by differentiating both sides of (16.8.12) with respect to T, holding p constant.

14.[R] One might try to find $(\partial v/\partial T)_p$ in Example 5 by first finding an equation that expresses v in terms of T and p. What unpleasantness happens when you try this approach?

15.[R] In Example 5, find $(\partial v/\partial p)_T$, $(\partial T/\partial v)_p$, and $(\partial T/\partial p)_v$.

16.[M] In thermodynamics there is the Gibbs relation

$$dh = T\ ds + v\ dP.$$

It is understood that $\left.\frac{\partial h}{\partial s}\right|_p = T$ and $\left.\frac{\partial h}{\partial p}\right|_s = v$. Deduce that $\left.\frac{\partial T}{\partial P}\right|_s = \left.\frac{\partial v}{\partial s}\right|_P$.

17.[R] Consider the thermodynamic equation

$$\left.\frac{\partial E}{\partial T}\right|_v = \left.\frac{\partial E}{\partial T}\right|_P + \left.\frac{\partial E}{\partial P}\right|_T \left.\frac{\partial P}{\partial T}\right|_v. \qquad (16.8.15)$$

(a) What is the dependent variable?

(b) What are the independent variables?

(c) What are the intermediate variables?

(d) Draw a diagram showing all the paths from the dependent variables to the independent variables.

(e) Use the Chain Rule to complete the derivation of (16.8.15).

18.[M] Show that $\left.\frac{\partial P}{\partial T}\right|_v = \frac{-\left.\frac{\partial v}{\partial T}\right|_P}{\left.\frac{\partial v}{\partial P}\right|_T}$.

19.[M] Show that

(a) $\left.\frac{\partial E}{\partial v}\right|_P = \left.\frac{\partial E}{\partial T}\right|_P \left.\frac{\partial T}{\partial v}\right|_P$

(b) $\left.\frac{\partial E}{\partial P}\right|_v = \left.\frac{\partial E}{\partial T}\right|_P \left.\frac{\partial T}{\partial P}\right|_v + \left.\frac{\partial E}{\partial P}\right|_T$.

20.[M] Show that $\left.\frac{\partial P}{\partial T}\right|_v \left.\frac{\partial T}{\partial P}\right|_v = 1$. HINT: Express each of the partial derivaties as a quotient of partial derivatives, as in Exercise 18.

21.[M] Show that $\left.\frac{\partial P}{\partial T}\right|_v \left.\frac{\partial T}{\partial v}\right|_P \left.\frac{\partial v}{\partial P}\right|_T = -1$.

22.[M] Let $u = F(x, y, z)$ and $z = f(x, y)$. Thus u is a (composite) function of x and y: $u = G(x, y) = F(x, y, f(x, y))$. Assume that $G(x, y) = x^2y$. Obtain a formula for $\frac{\partial f}{\partial x}$ in terms of $\frac{\partial F}{\partial x}$, $\frac{\partial F}{\partial y}$, and $\frac{\partial F}{\partial z}$. (All three need not appear in your answer.)

23.[M] Let $u = F(x, y, z)$ and $x = f(y, z)$. Thus u is a (composite) function of y and z: $u = G(y, z) = F(f(y, z), y, z)$. Assume that $G(y, z) = 2y + z^2$. Obtain a formula for $\frac{\partial f}{\partial z}$ in terms of $\frac{\partial F}{\partial x}$, $\frac{\partial F}{\partial y}$, and $\frac{\partial F}{\partial z}$. (All three need not appear in your answer.)

24.[C] Two functions u and v of the variables x and y are defined implicitly by the two simultaneous equations

$$F(u, v, x, y) = 0 \quad \text{and} \quad G(u, v, x, y) = 0.$$

Assuming all necessary differentiability, find a formula for $\frac{\partial u}{\partial x}$ in terms of the partial derivatives of F and of G.

16.S Chapter Summary

This chapter extends to functions of two or more variables the notions of rate of change and derivative originally in Chapter 3. For a function of several variables a "partial derivative" is simply the derivative with respect to one of the variables, when all the other variables are held constant.

The precise definition rests on a limit. For instance, the partial derivative with respect to x of $f(x, y)$ at (a, b) is

$$\frac{\partial f}{\partial x}(a,b) = \lim_{\Delta x \to 0} \frac{f(a+\Delta x, b) - f(a,b)}{\Delta x}.$$

Just as there are higher-order derivatives, there are higher-order partial derivatives, for instance:

$$\frac{\partial^2 f}{\partial x \partial x} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right), \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right), \frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right), \text{ and } \frac{\partial^2 f}{\partial y \partial y} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right)$$

For functions usually encountered in applications, the two "mixed partialas," $\partial^2 f/\partial x \partial y$ and $\partial^2 f/\partial y \partial x$, are equal; we can therefore not worry about the order of the differentiation.

Also, for common functions "differentiation under the integral sign" is legal:

$$\text{if } g(y) = \int_a^b f(x,y)\ dx, \text{ then } \frac{dg}{dy} = \int_a^b \frac{\partial f}{\partial y}(x,y)\ dx.$$

For a function of one variable, $f(x)$, with a continuous derivative,

$$\Delta f = f(a+\Delta x) - f(a) = f'(c)\Delta x = (f'(a)+\epsilon)\Delta x = f'(a)\Delta x + \epsilon \Delta x. \quad (16.S.1)$$

Here c is in $[a, a+\Delta x]$ and $\epsilon \to 0$ as $\Delta x \to 0$. The analog of (16.S.1) for a function of two or more variables is the basis for the chain rule for functions of several variables:

$$\Delta f = f(a+\Delta x, b+\Delta y) - f(a,b) = (f(a+\Delta x, b+\Delta y) - f(a, b+\Delta y)) + (f(a, b+\Delta y) \quad (16.S.2)$$

See the CIE section on Maxwell's equations at the end of Chapter 18.

where ϵ_1 and $\epsilon_2 \to 0$ as $Deltax$ and $\Delta y \to 0$.

The chain rule showed then, if $g(u)$ and $h(u)$ are differentiable functions, then $y = g(x+kt) + h(x-kt)$, k constant, satisfies the partial differential equation (PDE) $\partial^2 y/\partial t^2 = k^2 \partial^2 y/\partial x^2$. This PDE was the key to Maxwell's conjecture that light is an electro-magnetic phenomenon.

The gradient, a vector function, was defined in terms of partial derivatives: $\nabla f = \langle f_x, f_y \rangle$ or, for a function of three variables: $\nabla f = \langle f_x, f_y, f_z \rangle$. The gradient points in the direction a function increases most rapidly. The rate at which $f(x, y)$ changes in the direction of a unit vector $\mathbf{u}$ is $\nabla f \cdot \mathbf{u}$. The gradient is perpendicular to the level curve (or level surface) passing through a given point. At a critical point the gradient vanishes.

For a function of one variable the sign of the second derivative helps tell whether a critical point is a maximum or a minimum. For a function of two variables, the test also involves all three second derivatives. In particular, the signs of f_{xx} and $f_{xx}f_{yy} - (f_{xy})^2$ are important.

The number λ is called a Lagrange multiplier.

Maximizing a function f subject to a constraint g depends on the observation that at an extremum ∇f is parallel to ∇g. Hence there is a number λ such that $\nabla f = \lambda \nabla g$.

The final section showed that the chain rule is the bases of two facts in thermodynamics. It also shows how to apply the chain rule when a middle variable is also a final variable.

EXERCISES for 16.S *Key:* R–routine, M–moderate, C–challenging

1.[R] Let $f(x, y) = x^2 - y^2$ and $g(x, y) = 2xy$. Show that

(a) $\frac{\partial f}{\partial x} = \frac{\partial g}{\partial y}$

(b) $\frac{\partial f}{\partial y} = -\frac{\partial g}{\partial x}$

(c) $\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = 0$

(d) $\frac{\partial^2 g}{\partial x^2} + \frac{\partial^2 g}{\partial y^2} = 0$

2.[R] Repeat Exercise 1 for $f(x, y) = \ln\left(\sqrt{x^2 + y^2}\right)$ and $g(x, y) = \arctan(y/x)$.

3.[M] Let f and g be functions of x and y that have continuous second derivatives. Assume the first partial derivatives of f and g satisfy:

$$\frac{\partial f}{\partial x} = \frac{\partial g}{\partial y} \quad \text{and} \quad \frac{\partial f}{\partial y} = -\frac{\partial g}{\partial x}. \qquad (16.S.3)$$

Show that

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = 0 \quad \text{and} \quad \frac{\partial^2 g}{\partial x^2} + \frac{\partial^2 g}{\partial y^2} = 0. \qquad (16.S.4)$$

NOTE: The two equations in (16.S.3) are known as the Cauchy–Riemann equations. A pair of functions that satisfy (16.S.4) are called a **conformal pair** of functions.

In Exercises 4 to 12 assume the functions have continuous partial derivatives throughout the xy plane.

4.[R] If $f_x(x, y) = 0$ for all points (x, y) in the plane, must f be constant? If not, describe f.

5.[R] If $f_x(x, y) = 0$ and $f_y(x, y) = 0$ for all points (x, y) in the plane, must f be constant? If not, describe f.

6.[R] The function $3x + g(y)$, for any differentiable function $g(y)$ satisfies the partial differential equation $\partial f/\partial x = 3$. Are there any other solutions to that equation? Explain your answer.

7.[R] Find all functions f such that $\partial f/\partial x = 3$ and also $\partial y/\partial x = 3$ are satisfied.

8.[R] Show that there is no function f such that $\partial f/\partial x = 3y$ and $\partial f/\partial y = 4x$.

9.[R] Find all functions such that $f_{xx}(x, y) = 0$.

10.[R] Find all functions such that $f_{xx}(x, y) = 0$ and $f_{yy}(x, y) = 0$.

11.[R] Find all functions such that $f_{xy}(x, y) = 0$.

12.[R] Find all functions such that $f_{xy}(x, y) = 1$.

13.[M] A hiker is at the origin on a hill whose equation is $z = x$. If he walks south, along the positive x-axis the slope of his path would be steep, 1, with angle $\pi/4$. If he walked along the y-axis, the slope would be 0.

(a) If he walked NE what would the slope of his path be?

(b) In what direction should he walk in order that his path would have a slope of 0.2?

14.[C] This exercise outlines a proof that the two mixed partials of $f(x, y)$ are generally equal. It suffices to show that $f_{xy}(0, 0) = f_{yx}(0, 0)$. We assume that all the first and second partial derivatives are continuous in some disk with center $(0, 0)$.

(a) Why is $f_{xy}(0, 0)$ equal to

$$\lim_{k\to 0} \frac{f_x(0, k) - f_x(0, 0)}{k}? \tag{16.S.5}$$

(b) Why is (16.S.5) equal to

$$\lim_{k\to 0}\left(\lim_{h\to 0} \frac{(f(h, k) - f(0, k)) - (f(h, 0) - f(0, 0))}{hk}\right)? \tag{16.S.6}$$

(c) Let $u(y) = f(h, y) - f(0, y)$. Show that the fraction in (16.S.6) equals

$$\frac{u(k) - u(0)}{hk},$$

and this fraction equals $u'(k)/h$ for some k between 0 and k.

(d) Why is $u'(k) = f_y(h, k) - f_y(0, k)$?

(e) Why is $u'(k)/h$ equal to $(f_y)_x(H, K)$ for some H between 0 and h?

(f) Deduce that $f_{xy}(0, 0) = f_{yx}(0, 0)$.

(g) Did this derivation use the continuity of f_{yx}? If so, how?

(h) Did this derivation use the continuity of f_{xy}? If so, how?

(i) Did we need to assume f_{xy} exists? If so, where was this assumption used?

(j) Did we need to assume f_{yx} exists? If so, where was this assumption used?

15.[C] The assertion that it is safe to "differentiate across the integral sign," amounts to the statement that two definite integrals are equal. To illustrate this, translate the assertion into the language of limits:

$$\frac{d}{dt}\int_a^b f(x, t)\ dx = \int_a^b \frac{\partial}{\partial t} f(x, t)\ dx. \tag{16.S.7}$$

(a) Why is the derivative on the left an ordinary derivative, $d()/dt$, but the derivative on the right is a partial derivative?

(b) Using the definitions of ordinary derivatives and partial derivatives as limits, show what (16.S.7) says about limits.

(c) Verify (16.S.7) for $f(x, t) = x^7 t^4$.

(d) Verfiy (16.S.7) for $f(x, t) = \cos(xt)$.

Exercise 16 provides another motivation for the definition of the Fourier series of a function f defined on the interval $[0, 2\pi]$.

16.[C] For a particular integer n consider all functions

$S(x)$ of the form

$$S(x) = \frac{a_0}{2} + \sum_{k=1}^{n} \left(a_k \cos(kx) + b_k \sin(kx)\right),$$

Let $f(x)$ be a continuous function defined on $[0, 2\pi]$. The definite integral

$$\int_0^{2\pi} \left(f(x) - S(x)\right)^2 \, dx$$

is a measure of how close $S(x)$ is to $f(x)$ on the interval $[0, 2\pi]$. The integral can never be negative. (Why?) The smaller the integral, the better S approximates f on $[0, 2\pi]$. Show that the $S(x)$ that minimizes the integral is precisely a front-end of the Fourier series associated with $f(x)$.

Calculus is Everywhere # 21

The Wave in a Rope

We will develop what may be the most famous partial differential equation. In the CIE of the next chapter we will solve that equation and, then, use it in the final chapter to show how it helped Maxwell discover that light is an electrical-magnetic phenomenon.

As Morris Kline writes in *Mathematical Thought from Ancient to Modern Times*, "The first real success with partial differential equations came in renewed attacks on the vibrating string problem, typified by the violin string. The approximation that the vibrations are small was imposed by d'Alembert (1717-1783) in his papers of 1746."

Figure C.21.1:

Imagine shaking the end of a rope up and down gently, as in Figure C.21.1.

That motion starts a wave moving along the rope. The individual molecules in the rope move up and down, while the wave travels to the right. In the case of a sound wave, the wave travels at 700 miles per hour, but the air just vibrates back and forth. (When someone says "good morning" to us, we are not struck with a hurricane blast of wind.)

To develop the mathematics of the wave in a weightless rope, we begin with some simplifying assumptions. First, each molecule moves only up and down. Second, the distance each one moves is very small and the slope of the curve assumed by the rope remains close to zero. (Think of a violin string.)

At time t the vertical position of the molecule whose x-coordinate is x is $y = y(x, t)$, for it depends on both x and t. Consider a very short section of the rope at time t, shown as PQ in Figure C.21.2.

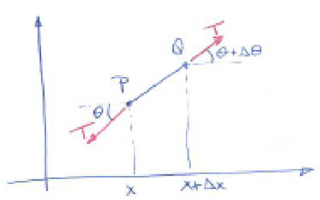


Figure C.21.2:

We assume that the tension T is the same throughout the rope. Apply Newton's Second Law, "force equals mass times acceleration," to the mass in PQ.

If the linear density of the rope is λ, the mass of the segment is λ times the length of the segment. Because we are assuming small displacements, we will approximate that length by Δx. The upward force exerted by the rope on the segment is $T\sin(\theta + \Delta\theta)$ and the downward force is $T\sin(\theta)$. The net vertical force is $T\sin(\theta + \Delta\theta) - T\sin(\theta)$. Thus

$$\underbrace{T sin(\theta + \Delta\theta) - T\sin(\theta)}_{\text{net vertical force}} = \underbrace{\lambda\Delta x}_{\text{mass}} \underbrace{\frac{\partial^2 y}{\partial t^2}}_{\text{acceleration}} . \tag{C.21.1}$$

(Because y is a function of x and t, we have a partial derivative, not an ordinary derivative.)

Next we express $\sin(\theta)$ and $\sin(\theta + \Delta\theta)$ in terms of the partial derivative $\partial y/\partial x$.

First of all, because θ is near 0, $\cos(\theta)$ is near 1. Thus $\sin(\theta)$ is approximately $\sin(\theta)/\cos(\theta) = \tan(\theta)$, the slope of the rope at time t above (or below) x, which is $\partial y/\partial x$ at x and t. Similarly, $\sin(\theta + \Delta\theta)$ is approximately $\partial y/\partial x$ at $x + \Delta x$ and t. So (C.21.1) is approximated by

$$T\frac{\partial y}{\partial x}(x + \Delta x, t) - T\frac{\partial y}{\partial x}(x, t) = \lambda \Delta x \frac{\partial^2 y}{\partial t^2}(x, t). \qquad \text{(C.21.2)}$$

Dividing both sides of (C.21.2) by Δx gives

$$\frac{T\left(\frac{\partial y}{\partial x}(x + \Delta x, t) - \frac{\partial y}{\partial x}(x, t)\right)}{\Delta x} = \lambda \frac{\partial^2 y}{\partial t^2}(x, t). \qquad \text{(C.21.3)}$$

Letting Δx in (C.21.3) approach 0, we obtain

$$T\frac{\partial^2 y}{\partial x^2}(x, t) = \lambda \frac{\partial^2 y}{\partial t^2}(x, t). \qquad \text{(C.21.4)}$$

Since both T and λ are positive, we can write (C.21.4) in the form

$$\frac{\partial^2 y}{\partial x^2} = \frac{1}{c^2}\frac{\partial^2 y}{\partial t^2}. \qquad \text{(C.21.5)}$$

This is the famous **wave equation**. It relates the acceleration of the molecule to the geometry of the curve; the latter is expressed by $\partial^2 y/\partial x^2$. Since we are assuming that the slope of the rope remains near 0, $\frac{\partial^2 y}{\partial x^2}$ is approximately

$$\frac{\frac{\partial^2 y}{\partial x^2}}{\left(\sqrt{1 + \left(\frac{\partial y}{\partial x}\right)^2}\right)^3}$$

which is the curvature at a given location and time. At the curvier part of the rope, the acceleration is greater.

As the CIE in the next chapter shows, the constant c turns out to be the velocity of the wave.

EXERCISES

1.[M] Figure C.21.3 shows a vibrating string whose ends are fixed at A and B. Assume that each part of the string moves parallel to the y-axis (a reasonable approximation of the vibrations are small.) Let $y = f(x, t)$ be the height of the string at the point with abscissa x at time t, as shown in the figure. In this case, the partial derivatives are denoted $\partial y/\partial x$ and $\partial y/\partial t$.

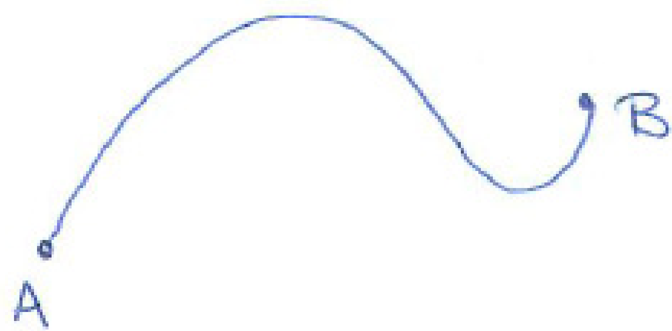


Figure C.21.3:

(a) What is the meaning of y_x?

(b) What is the meaning of y_t?