Section P.1

Graphs and Models

- **Sketch the graph of an equation.**
- **Find the intercepts of a graph.**
- **Test a graph for symmetry with respect to an axis and the origin.**
- **Find the points of intersection of two graphs.**
- **Interpret mathematical models for real-life data.**

RENÉ DESCARTES (1596–1650)
Descartes made many contributions to philosophy, science, and mathematics. The idea of representing points in the plane by pairs of real numbers and representing curves in the plane by equations was described by Descartes in his book *La Géométrie*, published in 1637.

MathBio

The Graph of an Equation

In 1637 the French mathematician René Descartes revolutionized the study of mathematics by joining its two major fields—algebra and geometry. With Descartes's coordinate plane, geometric concepts could be formulated analytically and algebraic concepts could be viewed graphically. The power of this approach is such that within a century, much of calculus had been developed.

The same approach can be followed in your study of calculus. That is, by viewing calculus from multiple perspectives—*graphically*, *analytically*, and *numerically*—you will increase your understanding of core concepts.

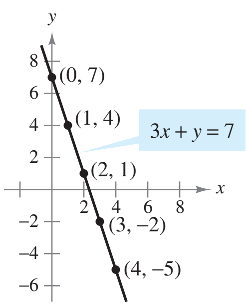
Consider the equation $3x + y = 7$. The point $(2, 1)$ is a **solution point** of the equation because the equation is satisfied (is true) when 2 is substituted for x and 1 is substituted for y. This equation has many other solutions, such as $(1, 4)$ and $(0, 7)$. To find other solutions systematically, solve the original equation for y.

$$y = 7 - 3x \qquad \text{Analytic approach}$$

Then construct a **table of values** by substituting several values of x.

x	0	1	2	3	4
y	7	4	1	-2	-5

Numerical approach



Graphical approach: $3x + y = 7$
Figure P.1

From the table, you can see that $(0, 7)$, $(1, 4)$, $(2, 1)$, $(3, -2)$, and $(4, -5)$ are solutions of the original equation $3x + y = 7$. Like many equations, this equation has an infinite number of solutions. The set of all solution points is the **graph** of the equation, as shown in Figure P.1.

NOTE Even though we refer to the sketch shown in Figure P.1 as the graph of $3x + y = 7$, it really represents only a *portion* of the graph. The entire graph would extend beyond the page.

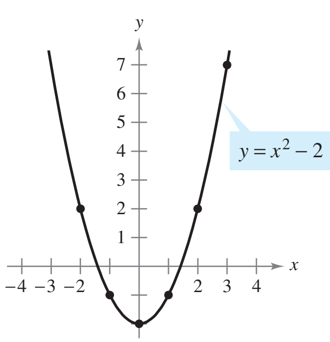
In this course, you will study many sketching techniques. The simplest is point plotting—that is, you plot points until the basic shape of the graph seems apparent.

EXAMPLE 1 Sketching a Graph by Point Plotting

Sketch the graph of $y = x^2 - 2$.

Solution First construct a table of values. Then plot the points shown in the table.

x	-2	-1	0	1	2	3
y	2	-1	-2	-1	2	7



The parabola $y = x^2 - 2$
Figure P.2

Finally, connect the points with a *smooth curve*, as shown in Figure P.2. This graph is a **parabola.** It is one of the conics you will study in Chapter 10.

Editable Graph

Try It

Exploration A

One disadvantage of point plotting is that to get a good idea about the shape of a graph, you may need to plot many points. With only a few points, you could badly misrepresent the graph. For instance, suppose that to sketch the graph of

$$y = \frac{1}{30}x(39 - 10x^2 + x^4)$$

you plotted only five points: $(-3, -3)$, $(-1, -1)$, $(0, 0)$, $(1, 1)$, and $(3, 3)$, as shown in Figure P.3(a). From these five points, you might conclude that the graph is a line. This, however, is not correct. By plotting several more points, you can see that the graph is more complicated, as shown in Figure P.3(b).

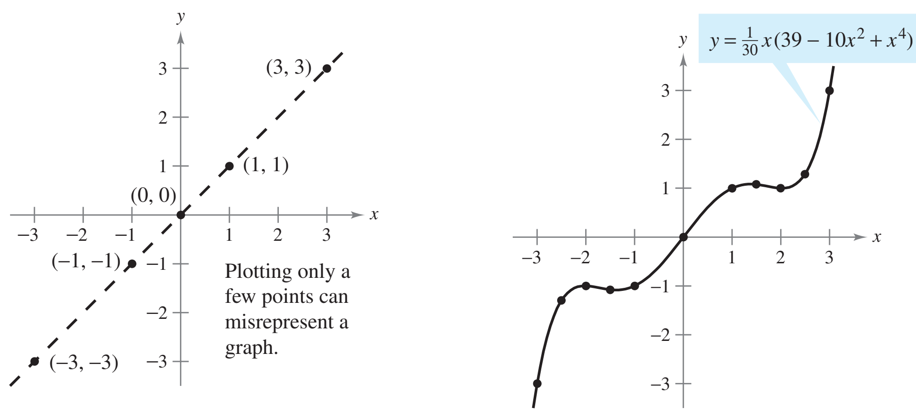


Figure P.3

EXPLORATION

Comparing Graphical and Analytic Approaches Use a graphing utility to graph each equation. In each case, find a viewing window that shows the important characteristics of the graph.

a. $y = x^3 - 3x^2 + 2x + 5$

b. $y = x^3 - 3x^2 + 2x + 25$

c. $y = -x^3 - 3x^2 + 20x + 5$

d. $y = 3x^3 - 40x^2 + 50x - 45$

e. $y = -(x + 12)^3$

f. $y = (x - 2)(x - 4)(x - 6)$

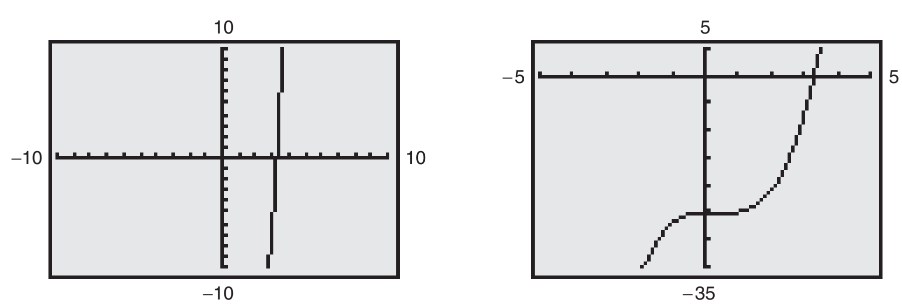
A purely graphical approach to this problem would involve a simple "guess, check, and revise" strategy. What types of things do you think an analytic approach might involve? For instance, does the graph have symmetry? Does the graph have turns? If so, where are they?

As you proceed through Chapters 1, 2, and 3 of this text, you will study many new analytic tools that will help you analyze graphs of equations such as these.

TECHNOLOGY Technology has made sketching of graphs easier. Even with technology, however, it is possible to misrepresent a graph badly. For instance, each of the graphing utility screens in Figure P.4 shows a portion of the graph of

$$y = x^3 - x^2 - 25.$$

From the screen on the left, you might assume that the graph is a line. From the screen on the right, however, you can see that the graph is not a line. So, whether you are sketching a graph by hand or using a graphing utility, you must realize that different "viewing windows" can produce very different views of a graph. In choosing a viewing window, your goal is to show a view of the graph that fits well in the context of the problem.



Graphing utility screens of $y = x^3 - x^2 - 25$
Figure P.4

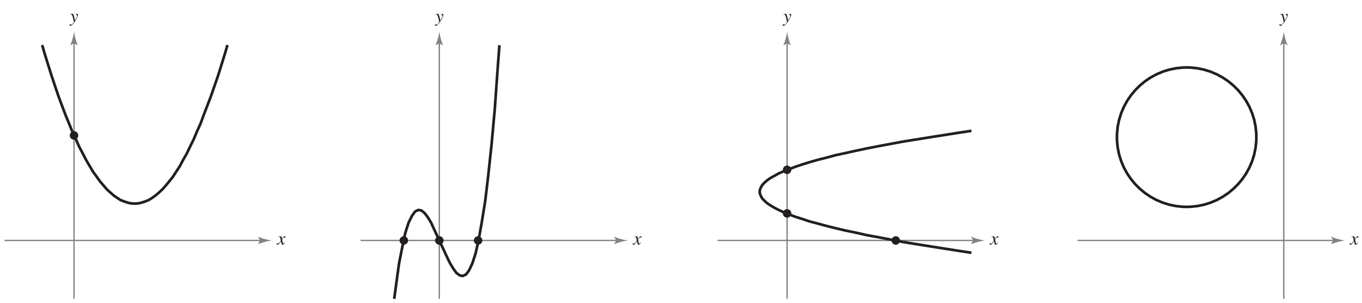
NOTE In this text, the term *graphing utility* means either a graphing calculator or computer graphing software such as *Maple*, *Mathematica*, *Derive*, *Mathcad*, or the *TI-89*.

Intercepts of a Graph

Two types of solution points that are especially useful in graphing an equation are those having zero as their x- or y-coordinate. Such points are called **intercepts** because they are the points at which the graph intersects the x- or y-axis. The point $(a, 0)$ is an ***x*-intercept** of the graph of an equation if it is a solution point of the equation. To find the x-intercepts of a graph, let y be zero and solve the equation for x. The point $(0, b)$ is a ***y*-intercept** of the graph of an equation if it is a solution point of the equation. To find the y-intercepts of a graph, let x be zero and solve the equation for y.

NOTE Some texts denote the x-intercept as the x-coordinate of the point $(a, 0)$ rather than the point itself. Unless it is necessary to make a distinction, we will use the term *intercept* to mean either the point or the coordinate.

It is possible for a graph to have no intercepts, or it might have several. For instance, consider the four graphs shown in Figure P.5.



No x-intercepts
One y-intercept

Three x-intercepts
One y-intercept

One x-intercept
Two y-intercepts

No intercepts

Figure P.5

EXAMPLE 2 Finding *x*- and *y*-intercepts

Find the x- and y-intercepts of the graph of $y = x^3 - 4x$.

Solution To find the x-intercepts, let y be zero and solve for x.

$$x^3 - 4x = 0 \qquad \text{Let } y \text{ be zero.}$$

$$x(x - 2)(x + 2) = 0 \qquad \text{Factor.}$$

$$x = 0, 2, \text{ or } -2 \qquad \text{Solve for } x.$$

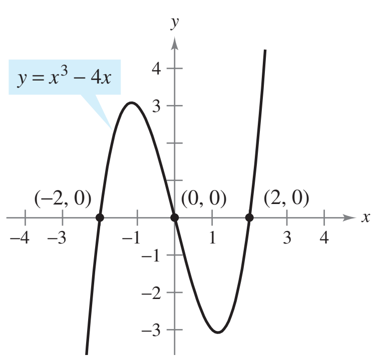
Because this equation has three solutions, you can conclude that the graph has three x-intercepts:

$$(0, 0), (2, 0), \text{ and } (-2, 0). \qquad x\text{-intercepts}$$

To find the y-intercepts, let x be zero. Doing this produces $y = 0$. So, the y-intercept is

$$(0, 0). \qquad y\text{-intercept}$$

(See Figure P.6.)



Intercepts of a graph
Figure P.6

Editable Graph

Try It

Exploration A

Video

Video

TECHNOLOGY Example 2 uses an analytic approach to finding intercepts. When an analytic approach is not possible, you can use a graphical approach by finding the points at which the graph intersects the axes. Use a graphing utility to approximate the intercepts.

Symmetry of a Graph

Knowing the symmetry of a graph *before* attempting to sketch it is useful because you need only half as many points to sketch the graph. The following three types of symmetry can be used to help sketch the graphs of equations (see Figure P.7).

1. A graph is **symmetric with respect to the *y*-axis** if, whenever (x, y) is a point on the graph, $(-x, y)$ is also a point on the graph. This means that the portion of the graph to the left of the y-axis is a mirror image of the portion to the right of the y-axis.
2. A graph is **symmetric with respect to the *x*-axis** if, whenever (x, y) is a point on the graph, $(x, -y)$ is also a point on the graph. This means that the portion of the graph above the x-axis is a mirror image of the portion below the x-axis.
3. A graph is **symmetric with respect to the origin** if, whenever (x, y) is a point on the graph, $(-x, -y)$ is also a point on the graph. This means that the graph is unchanged by a rotation of 180° about the origin.

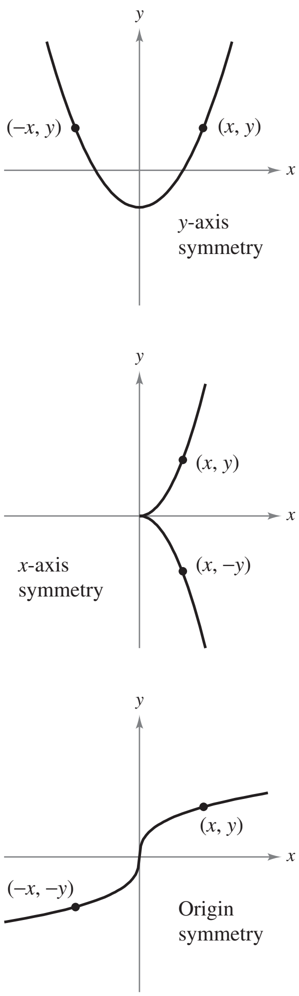


Figure P.7

Tests for Symmetry

1. The graph of an equation in x and y is symmetric with respect to the y-axis if replacing x by $-x$ yields an equivalent equation.
2. The graph of an equation in x and y is symmetric with respect to the x-axis if replacing y by $-y$ yields an equivalent equation.
3. The graph of an equation in x and y is symmetric with respect to the origin if replacing x by $-x$ and y by $-y$ yields an equivalent equation.

The graph of a polynomial has symmetry with respect to the y-axis if each term has an even exponent (or is a constant). For instance, the graph of

$$y = 2x^4 - x^2 + 2 \qquad \text{y-axis symmetry}$$

has symmetry with respect to the y-axis. Similarly, the graph of a polynomial has symmetry with respect to the origin if each term has an odd exponent, as illustrated in Example 3.

EXAMPLE 3 Testing for Origin Symmetry

Show that the graph of

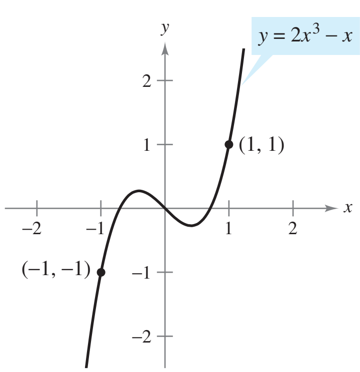
$$y = 2x^3 - x$$

is symmetric with respect to the origin.

Solution

$$\begin{aligned} y &= 2x^3 - x && \text{Write original equation.} \\ -y &= 2(-x)^3 - (-x) && \text{Replace x by $-x$ and y by $-y$.} \\ -y &= -2x^3 + x && \text{Simplify.} \\ y &= 2x^3 - x && \text{Equivalent equation} \end{aligned}$$

Because the replacements yield an equivalent equation, you can conclude that the graph of $y = 2x^3 - x$ is symmetric with respect to the origin, as shown in Figure P.8.



Origin symmetry
Figure P.8

Editable Graph

Try It Exploration A Video Video

EXAMPLE 4 Using Intercepts and Symmetry to Sketch a Graph

Sketch the graph of $x - y^2 = 1$.

Solution The graph is symmetric with respect to the x-axis because replacing y by $-y$ yields an equivalent equation.

$x - y^2 = 1$	Write original equation.
$x - (-y)^2 = 1$	Replace y by $-y$.
$x - y^2 = 1$	Equivalent equation

This means that the portion of the graph below the x-axis is a mirror image of the portion above the x-axis. To sketch the graph, first plot the x-intercept and the points above the x-axis. Then reflect in the x-axis to obtain the entire graph, as shown in Figure P.9.

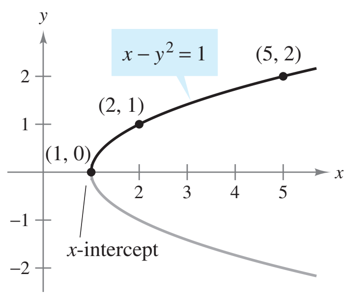


Figure P.9

Editable Graph

Try It | Exploration A | Exploration B | Open Exploration

TECHNOLOGY Graphing utilities are designed so that they most easily graph equations in which y is a function of x (see Section P.3 for a definition of **function**). To graph other types of equations, you need to split the graph into two or more parts *or* you need to use a different graphing mode. For instance, to graph the equation in Example 4, you can split it into two parts.

$y_1 = \sqrt{x - 1}$	Top portion of graph
$y_2 = -\sqrt{x - 1}$	Bottom portion of graph

Points of Intersection

A **point of intersection** of the graphs of two equations is a point that satisfies both equations. You can find the points of intersection of two graphs by solving their equations simultaneously.

EXAMPLE 5 Finding Points of Intersection

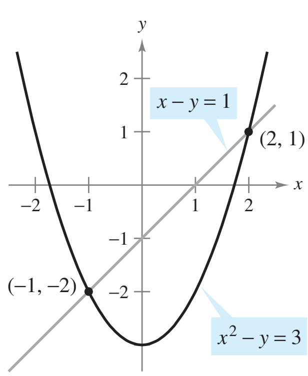
Find all points of intersection of the graphs of $x^2 - y = 3$ and $x - y = 1$.

Solution Begin by sketching the graphs of both equations on the *same* rectangular coordinate system, as shown in Figure P.10. Having done this, it appears that the graphs have two points of intersection. You can find these two points, as follows.

$y = x^2 - 3$	Solve first equation for y.
$y = x - 1$	Solve second equation for y.
$x^2 - 3 = x - 1$	Equate y-values.
$x^2 - x - 2 = 0$	Write in general form.
$(x - 2)(x + 1) = 0$	Factor.
$x = 2 \text{ or } -1$	Solve for x.

The corresponding values of y are obtained by substituting $x = 2$ and $x = -1$ into either of the original equations. Doing this produces two points of intersection:

$(2, 1)$ and $(-1, -2)$. Points of intersection



Two points of intersection
Figure P.10

Editable Graph

STUDY TIP You can check the points of intersection from Example 5 by substituting into *both* of the original equations or by using the *intersect* feature of a graphing utility.

Try It | Exploration A

Mathematical Models

Real-life applications of mathematics often use equations as **mathematical models.** In developing a mathematical model to represent actual data, you should strive for two (often conflicting) goals: accuracy and simplicity. That is, you want the model to be simple enough to be workable, yet accurate enough to produce meaningful results. Section P.4 explores these goals more completely.

EXAMPLE 6 Comparing Two Mathematical Models

The Mauna Loa Observatory in Hawaii has been measuring the increasing concentration of carbon dioxide in Earth's atmosphere since 1958.

Video

The Mauna Loa Observatory in Hawaii records the carbon dioxide concentration y (in parts per million) in Earth's atmosphere. The January readings for various years are shown in Figure P.11. In the July 1990 issue of *Scientific American*, these data were used to predict the carbon dioxide level in Earth's atmosphere in the year 2035, using the quadratic model

$$y = 316.2 + 0.70t + 0.018t^2 \qquad \text{Quadratic model for 1960–1990 data}$$

where $t = 0$ represents 1960, as shown in Figure P.11(a).

The data shown in Figure P.11(b) represent the years 1980 through 2002 and can be modeled by

$$y = 306.3 + 1.56t \qquad \text{Linear model for 1980–2002 data}$$

where $t = 0$ represents 1960. What was the prediction given in the *Scientific American* article in 1990? Given the new data for 1990 through 2002, does this prediction for the year 2035 seem accurate?

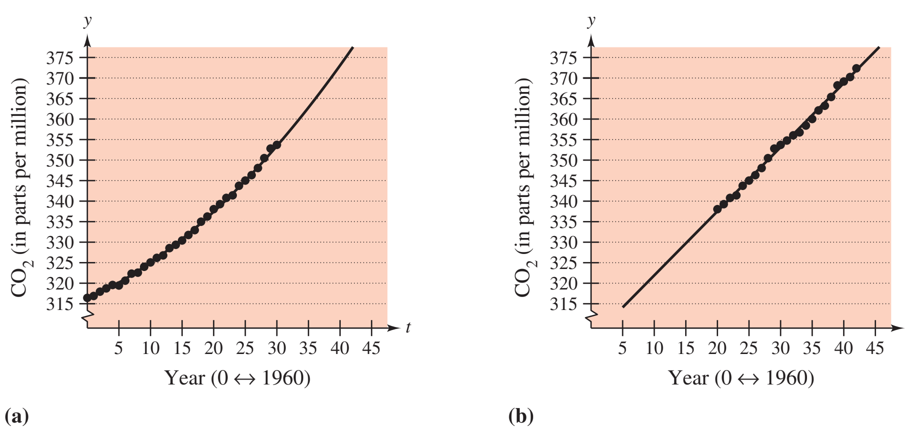


Figure P.11

Solution To answer the first question, substitute $t = 75$ (for 2035) into the quadratic model.

$$y = 316.2 + 0.70(75) + 0.018(75)^2 = 469.95 \qquad \text{Quadratic model}$$

So, the prediction in the *Scientific American* article was that the carbon dioxide concentration in Earth's atmosphere would reach about 470 parts per million in the year 2035. Using the linear model for the 1980–2002 data, the prediction for the year 2035 is

$$y = 306.3 + 1.56(75) = 423.3. \qquad \text{Linear model}$$

So, based on the linear model for 1980–2002, it appears that the 1990 prediction was too high.

NOTE The models in Example 6 were developed using a procedure called *least squares regression* (see Section 13.9). The quadratic and linear models have a correlation given by $r^2 = 0.997$ and $r^2 = 0.996$, respectively. The closer r^2 is to 1, the "better" the model.

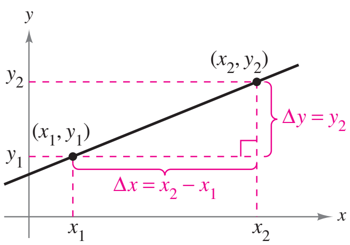
Try It

Exploration A

Section P.2 Linear Models and Rates of Change

- **Find the slope of a line passing through two points.**
- **Write the equation of a line with a given point and slope.**
- **Interpret slope as a ratio or as a rate in a real-life application.**
- **Sketch the graph of a linear equation in slope-intercept form.**
- **Write equations of lines that are parallel or perpendicular to a given line.**

The Slope of a Line



$\Delta y = y_2 - y_1 =$ change in y
$\Delta x = x_2 - x_1 =$ change in x
Figure P.12

The **slope** of a nonvertical line is a measure of the number of units the line rises (or falls) vertically for each unit of horizontal change from left to right. Consider the two points (x_1, y_1) and (x_2, y_2) on the line in Figure P.12. As you move from left to right along this line, a vertical change of

$$\Delta y = y_2 - y_1 \qquad \text{Change in } y$$

units corresponds to a horizontal change of

$$\Delta x = x_2 - x_1 \qquad \text{Change in } x$$

units. (Δ is the Greek uppercase letter *delta*, and the symbols Δy and Δx are read "delta y" and "delta x.")

Definition of the Slope of a Line

The **slope** m of the nonvertical line passing through (x_1, y_1) and (x_2, y_2) is

$$m = \frac{\Delta y}{\Delta x} = \frac{y_2 - y_1}{x_2 - x_1}, \qquad x_1 \neq x_2.$$

Slope is not defined for vertical lines.

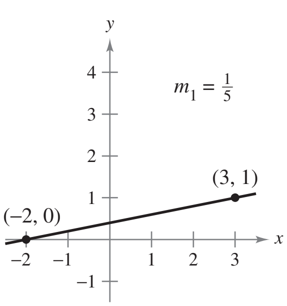
Video Video Video

NOTE When using the formula for slope, note that

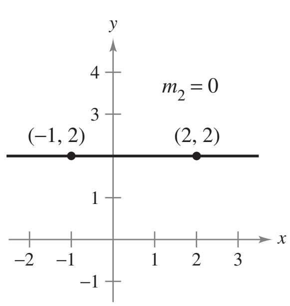
$$\frac{y_2 - y_1}{x_2 - x_1} = \frac{-(y_1 - y_2)}{-(x_1 - x_2)} = \frac{y_1 - y_2}{x_1 - x_2}.$$

So, it does not matter in which order you subtract *as long as* you are consistent and both "subtracted coordinates" come from the same point.

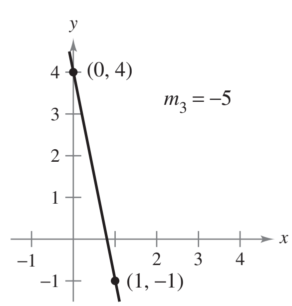
Figure P.13 shows four lines: one has a positive slope, one has a slope of zero, one has a negative slope, and one has an "undefined" slope. In general, the greater the absolute value of the slope of a line, the steeper the line is. For instance, in Figure P.13, the line with a slope of -5 is steeper than the line with a slope of $\frac{1}{5}$.



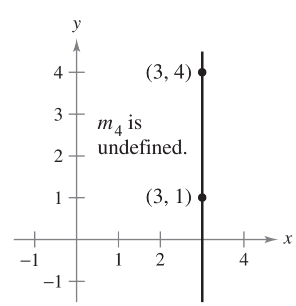
If m is positive, then the line rises from left to right.



If m is zero, then the line is horizontal.



If m is negative, then the line falls from left to right.

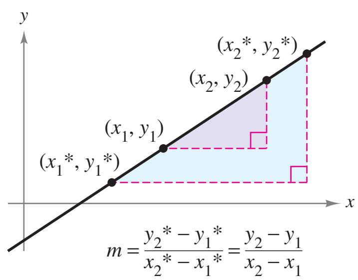


If m is undefined, then the line is vertical.

Figure P.13

Equations of Lines

Any two points on a nonvertical line can be used to calculate its slope. This can be verified from the similar triangles shown in Figure P.14. (Recall that the ratios of corresponding sides of similar triangles are equal.)



Any two points on a nonvertical line can be used to determine its slope.
Figure P.14

EXPLORATION

Investigating Equations of Lines Use a graphing utility to graph each of the linear equations. Which point is common to all seven lines? Which value in the equation determines the slope of each line?

a. $y - 4 = -2(x + 1)$
b. $y - 4 = -1(x + 1)$
c. $y - 4 = -\frac{1}{2}(x + 1)$
d. $y - 4 = 0(x + 1)$
e. $y - 4 = \frac{1}{2}(x + 1)$
f. $y - 4 = 1(x + 1)$
g. $y - 4 = 2(x + 1)$

Use your results to write an equation of a line passing through $(-1, 4)$ with a slope of m.

You can write an equation of a nonvertical line if you know the slope of the line and the coordinates of one point on the line. Suppose the slope is m and the point is (x_1, y_1). If (x, y) is any other point on the line, then

$$\frac{y - y_1}{x - x_1} = m.$$

This equation, involving the two variables x and y, can be rewritten in the form $y - y_1 = m(x - x_1)$, which is called the **point-slope equation of a line.**

Point-Slope Equation of a Line

An equation of the line with slope m passing through the point (x_1, y_1) is given by

$$y - y_1 = m(x - x_1).$$

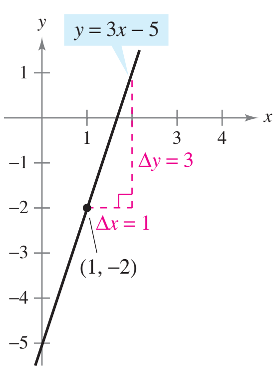
EXAMPLE 1 Finding an Equation of a Line

Find an equation of the line that has a slope of 3 and passes through the point $(1, -2)$.

Solution

$y - y_1 = m(x - x_1)$	Point-slope form
$y - (-2) = 3(x - 1)$	Substitute -2 for y_1, 1 for x_1, and 3 for m.
$y + 2 = 3x - 3$	Simplify.
$y = 3x - 5$	Solve for y.

(See Figure P.15.)



The line with a slope of 3 passing through the point $(1, -2)$
Figure P.15

Editable Graph

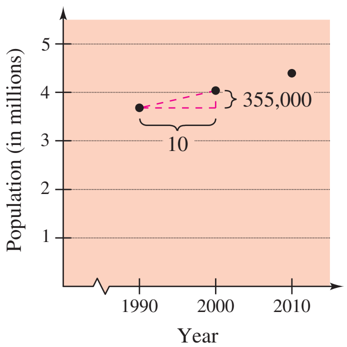
Try It   Exploration A   Exploration B   Exploration C

NOTE Remember that only nonvertical lines have a slope. Consequently, vertical lines cannot be written in point-slope form. For instance, the equation of the vertical line passing through the point $(1, -2)$ is $x = 1$.

Ratios and Rates of Change

The slope of a line can be interpreted as either a *ratio* or a *rate*. If the x- and y-axes have the same unit of measure, the slope has no units and is a **ratio.** If the x- and y-axes have different units of measure, the slope is a rate or **rate of change.** In your study of calculus, you will encounter applications involving both interpretations of slope.

EXAMPLE 2 Population Growth and Engineering Design



Population of Kentucky in census years
Figure P.16

a. The population of Kentucky was 3,687,000 in 1990 and 4,042,000 in 2000. Over this 10-year period, the average rate of change of the population was

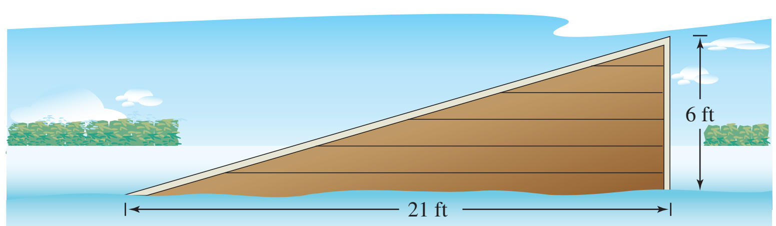
$$\begin{aligned}\text{Rate of change} &= \frac{\text{change in population}}{\text{change in years}}\\ &= \frac{4{,}042{,}000 - 3{,}687{,}000}{2000 - 1990}\\ &= 35{,}500 \text{ people per year.}\end{aligned}$$

If Kentucky's population continues to increase at this same rate for the next 10 years, it will have a 2010 population of 4,397,000 (see Figure P.16). *(Source: U.S. Census Bureau)*

b. In tournament water-ski jumping, the ramp rises to a height of 6 feet on a raft that is 21 feet long, as shown in Figure P.17. The slope of the ski ramp is the ratio of its height (the rise) to the length of its base (the run).

$$\begin{aligned}\text{Slope of ramp} &= \frac{\text{rise}}{\text{run}} && \text{Rise is vertical change, run is horizontal change.}\\ &= \frac{6 \text{ feet}}{21 \text{ feet}}\\ &= \frac{2}{7}\end{aligned}$$

In this case, note that the slope is a ratio and has no units.



Dimensions of a water-ski ramp
Figure P.17

Try It | Exploration A | Exploration B

The rate of change found in Example 2(a) is an **average rate of change.** An average rate of change is always calculated over an interval. In this case, the interval is [1990, 2000]. In Chapter 2 you will study another type of rate of change called an *instantaneous rate of change*.

Graphing Linear Models

Many problems in analytic geometry can be classified in two basic categories: (1) Given a graph, what is its equation? and (2) Given an equation, what is its graph? The point-slope equation of a line can be used to solve problems in the first category. However, this form is not especially useful for solving problems in the second category. The form that is better suited to sketching the graph of a line is the **slope-intercept** form of the equation of a line.

The Slope-Intercept Equation of a Line

The graph of the linear equation

$$y = mx + b$$

is a line having a *slope* of m and a y-*intercept* at $(0, b)$.

Video Video

EXAMPLE 3 Sketching Lines in the Plane

Sketch the graph of each equation.

a. $y = 2x + 1$ **b.** $y = 2$ **c.** $3y + x - 6 = 0$

Solution

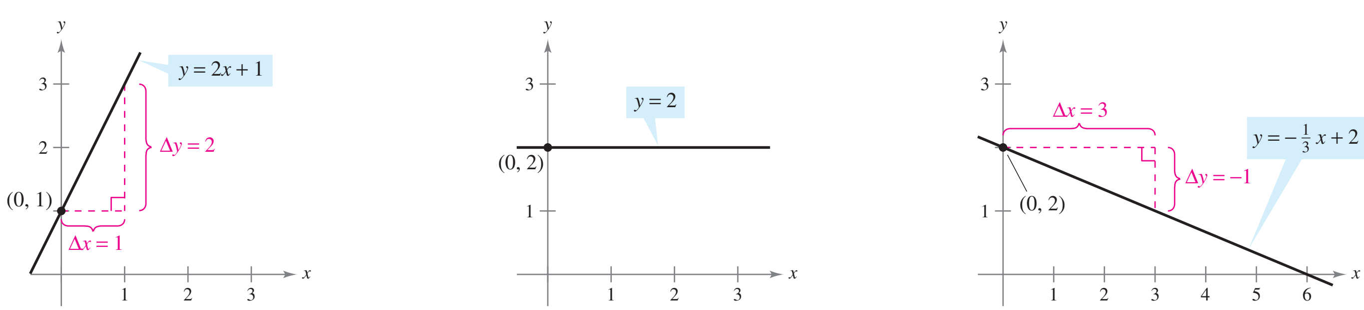
a. Because $b = 1$, the y-intercept is $(0, 1)$. Because the slope is $m = 2$, you know that the line rises two units for each unit it moves to the right, as shown in Figure P.18(a).

b. Because $b = 2$, the y-intercept is $(0, 2)$. Because the slope is $m = 0$, you know that the line is horizontal, as shown in Figure P.18(b).

c. Begin by writing the equation in slope-intercept form.

$$3y + x - 6 = 0 \qquad \text{Write original equation.}$$
$$3y = -x + 6 \qquad \text{Isolate } y\text{-term on the left.}$$
$$y = -\frac{1}{3}x + 2 \qquad \text{Slope-intercept form}$$

In this form, you can see that the y-intercept is $(0, 2)$ and the slope is $m = -\frac{1}{3}$. This means that the line falls one unit for every three units it moves to the right, as shown in Figure P.18(c).



(a) $m = 2$; line rises **(b)** $m = 0$; line is horizontal **(c)** $m = -\frac{1}{3}$; line falls

Figure P.18

Editable Graph Editable Graph Editable Graph

Try It Exploration A

Because the slope of a vertical line is not defined, its equation cannot be written in the slope-intercept form. However, the equation of any line can be written in the **general form**

$$Ax + By + C = 0$$

General form of the equation of a line

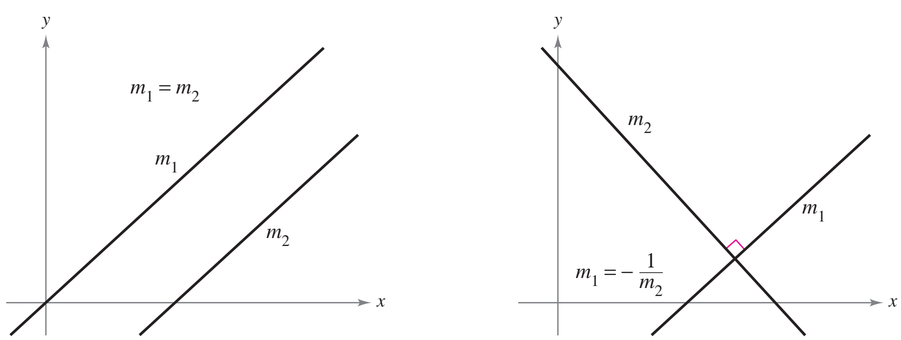
where A and B are not *both* zero. For instance, the vertical line given by $x = a$ can be represented by the general form $x - a = 0$.

Summary of Equations of Lines

1. General form: $Ax + By + C = 0, \quad (A, B \neq 0)$
2. Vertical line: $x = a$
3. Horizontal line: $y = b$
4. Point-slope form: $y - y_1 = m(x - x_1)$
5. Slope-intercept form: $y = mx + b$

Parallel and Perpendicular Lines

The slope of a line is a convenient tool for determining whether two lines are parallel or perpendicular, as shown in Figure P.19. Specifically, nonvertical lines with the same slope are parallel and nonvertical lines whose slopes are negative reciprocals are perpendicular.



Parallel lines
Perpendicular lines
Figure P.19

STUDY TIP In mathematics, the phrase "if and only if" is a way of stating two implications in one statement. For instance, the first statement at the right could be rewritten as the following two implications.

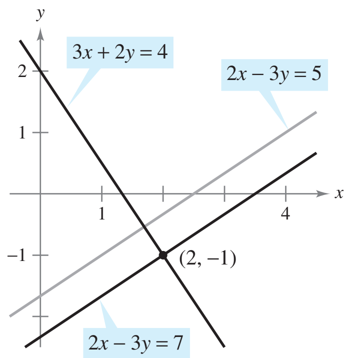
a. If two distinct nonvertical lines are parallel, then their slopes are equal.

b. If two distinct nonvertical lines have equal slopes, then they are parallel.

Parallel and Perpendicular Lines

1. Two distinct nonvertical lines are **parallel** if and only if their slopes are equal—that is, if and only if $m_1 = m_2$.
2. Two nonvertical lines are **perpendicular** if and only if their slopes are negative reciprocals of each other—that is, if and only if

$$m_1 = -\frac{1}{m_2}.$$



Lines parallel and perpendicular to $2x - 3y = 5$

Figure P.20

Editable Graph

EXAMPLE 4 Finding Parallel and Perpendicular Lines

Find the general forms of the equations of the lines that pass through the point $(2, -1)$ and are

a. parallel to the line $2x - 3y = 5$ **b.** perpendicular to the line $2x - 3y = 5$.

(See Figure P.20.)

Solution By writing the linear equation $2x - 3y = 5$ in slope-intercept form, $y = \frac{2}{3}x - \frac{5}{3}$, you can see that the given line has a slope of $m = \frac{2}{3}$.

a. The line through $(2, -1)$ that is parallel to the given line also has a slope of $\frac{2}{3}$.

$$
\begin{aligned}
y - y_1 &= m(x - x_1) && \text{Point-slope form}\\
y - (-1) &= \tfrac{2}{3}(x - 2) && \text{Substitute.}\\
3(y + 1) &= 2(x - 2) && \text{Simplify.}\\
2x - 3y - 7 &= 0 && \text{General form}
\end{aligned}
$$

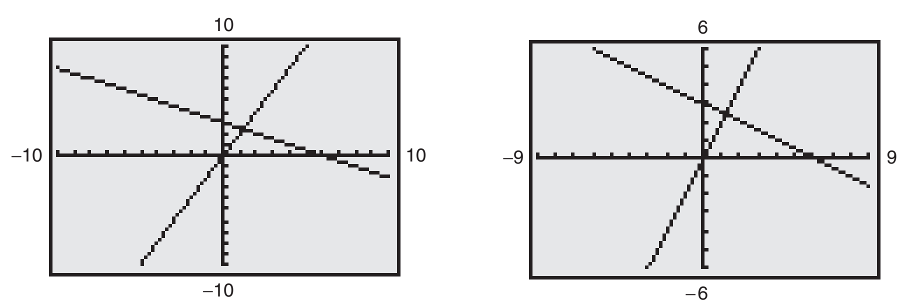
Note the similarity to the original equation.

b. Using the negative reciprocal of the slope of the given line, you can determine that the slope of a line perpendicular to the given line is $-\frac{3}{2}$. So, the line through the point $(2, -1)$ that is perpendicular to the given line has the following equation.

$$
\begin{aligned}
y - y_1 &= m(x - x_1) && \text{Point-slope form}\\
y - (-1) &= -\tfrac{3}{2}(x - 2) && \text{Substitute.}\\
2(y + 1) &= -3(x - 2) && \text{Simplify.}\\
3x + 2y - 4 &= 0 && \text{General form}
\end{aligned}
$$

Try It | Exploration A | Exploration B | Exploration C | Open Exploration

TECHNOLOGY PITFALL The slope of a line will appear distorted if you use different tick-mark spacing on the x- and y-axes. For instance, the graphing calculator screens in Figures P.21(a) and P.21(b) both show the lines given by $y = 2x$ and $y = -\frac{1}{2}x + 3$. Because these lines have slopes that are negative reciprocals, they must be perpendicular. In Figure P.21(a), however, the lines don't appear to be perpendicular because the tick-mark spacing on the x-axis is not the same as that on the y-axis. In Figure P.21(b), the lines appear perpendicular because the tick-mark spacing on the x-axis is the same as on the y-axis. This type of viewing window is said to have a *square setting*.



(a) Tick-mark spacing on the x-axis is not the same as tick-mark spacing on the y-axis.

(b) Tick-mark spacing on the x-axis is the same as tick-mark spacing on the y-axis.

Figure P.21

Section P.3 Functions and Their Graphs

- **Use function notation to represent and evaluate a function.**
- **Find the domain and range of a function.**
- **Sketch the graph of a function.**
- **Identify different types of transformations of functions.**
- **Classify functions and recognize combinations of functions.**

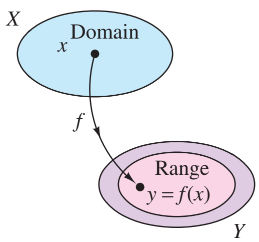
Functions and Function Notation

A **relation** between two sets X and Y is a set of ordered pairs, each of the form (x, y), where x is a member of X and y is a member of Y. A **function** from X to Y is a relation between X and Y that has the property that any two ordered pairs with the same x-value also have the same y-value. The variable x is the **independent variable,** and the variable y is the **dependent variable.**

Many real-life situations can be modeled by functions. For instance, the area A of a circle is a function of the circle's radius r.

$$A = \pi r^2 \qquad \text{A is a function of r.}$$

In this case r is the independent variable and A is the dependent variable.



A real-valued function f of a real variable
Figure P.22

Definition of a Real-Valued Function of a Real Variable

Let X and Y be sets of real numbers. A **real-valued function f of a real variable x** from X to Y is a correspondence that assigns to each number x in X exactly one number y in Y.

The **domain** of f is the set X. The number y is the **image** of x under f and is denoted by $f(x)$, which is called the **value of f at x.** The **range** of f is a subset of Y and consists of all images of numbers in X (see Figure P.22).

Functions can be specified in a variety of ways. In this text, however, we will concentrate primarily on functions that are given by equations involving the dependent and independent variables. For instance, the equation

$$x^2 + 2y = 1 \qquad \text{Equation in implicit form}$$

defines y, the dependent variable, as a function of x, the independent variable. To **evaluate** this function (that is, to find the y-value that corresponds to a given x-value), it is convenient to isolate y on the left side of the equation.

$$y = \frac{1}{2}(1 - x^2) \qquad \text{Equation in explicit form}$$

Using f as the name of the function, you can write this equation as

$$f(x) = \frac{1}{2}(1 - x^2). \qquad \text{Function notation}$$

The original equation, $x^2 + 2y = 1$, **implicitly** defines y as a function of x. When you solve the equation for y, you are writing the equation in **explicit** form.

Function notation has the advantage of clearly identifying the dependent variable as $f(x)$ while at the same time telling you that x is the independent variable and that the function itself is "f." The symbol $f(x)$ is read "f of x." Function notation allows you to be less wordy. Instead of asking "What is the value of y that corresponds to $x = 3$?" you can ask "What is $f(3)$?"

> **FUNCTION NOTATION**
>
> The word *function* was first used by Gottfried Wilhelm Leibniz in 1694 as a term to denote any quantity connected with a curve, such as the coordinates of a point on a curve or the slope of a curve. Forty years later, Leonhard Euler used the word "function" to describe any expression made up of a variable and some constants. He introduced the notation $y = f(x)$.

In an equation that defines a function, the role of the variable x is simply that of a placeholder. For instance, the function given by

$$f(x) = 2x^2 - 4x + 1$$

can be described by the form

$$f(\quad) = 2(\quad)^2 - 4(\quad) + 1$$

where parentheses are used instead of x. To evaluate $f(-2)$, simply place -2 in each set of parentheses.

$$\begin{aligned} f(-2) &= 2(-2)^2 - 4(-2) + 1 && \text{Substitute } -2 \text{ for } x. \\ &= 2(4) + 8 + 1 && \text{Simplify.} \\ &= 17 && \text{Simplify.} \end{aligned}$$

NOTE Although f is often used as a convenient function name and x as the independent variable, you can use other symbols. For instance, the following equations all define the same function.

$$\begin{aligned} f(x) &= x^2 - 4x + 7 && \text{Function name is } f\text{, independent variable is } x. \\ f(t) &= t^2 - 4t + 7 && \text{Function name is } f\text{, independent variable is } t. \\ g(s) &= s^2 - 4s + 7 && \text{Function name is } g\text{, independent variable is } s. \end{aligned}$$

EXAMPLE 1 Evaluating a Function

For the function f defined by $f(x) = x^2 + 7$, evaluate each expression.

a. $f(3a)$ **b.** $f(b - 1)$ **c.** $\dfrac{f(x + \Delta x) - f(x)}{\Delta x}, \quad \Delta x \neq 0$

Solution

$$\begin{aligned} \textbf{a. } f(3a) &= (3a)^2 + 7 && \text{Substitute } 3a \text{ for } x. \\ &= 9a^2 + 7 && \text{Simplify.} \end{aligned}$$

$$\begin{aligned} \textbf{b. } f(b - 1) &= (b - 1)^2 + 7 && \text{Substitute } b - 1 \text{ for } x. \\ &= b^2 - 2b + 1 + 7 && \text{Expand binomial.} \\ &= b^2 - 2b + 8 && \text{Simplify.} \end{aligned}$$

$$\begin{aligned} \textbf{c. } \frac{f(x + \Delta x) - f(x)}{\Delta x} &= \frac{[(x + \Delta x)^2 + 7] - (x^2 + 7)}{\Delta x} \\ &= \frac{x^2 + 2x\Delta x + (\Delta x)^2 + 7 - x^2 - 7}{\Delta x} \\ &= \frac{2x\Delta x + (\Delta x)^2}{\Delta x} \\ &= \frac{\cancel{\Delta x}(2x + \Delta x)}{\cancel{\Delta x}} \\ &= 2x + \Delta x, \quad \Delta x \neq 0 \end{aligned}$$

Try It Exploration A

NOTE The expression in Example 1(c) is called a *difference quotient* and has a special significance in calculus. You will learn more about this in Chapter 2.

STUDY TIP In calculus, it is important to communicate clearly the domain of a function or expression. For instance, in Example 1(c) the two expressions

$$\frac{f(x + \Delta x) - f(x)}{\Delta x} \quad \text{and} \quad 2x + \Delta x,$$
$\Delta x \neq 0$

are equivalent because $\Delta x = 0$ is excluded from the domain of each expression. Without a stated domain restriction, the two expressions would not be equivalent.

The Domain and Range of a Function

The domain of a function can be described explicitly, or it may be described *implicitly* by an equation used to define the function. The implied domain is the set of all real numbers for which the equation is defined, whereas an explicitly defined domain is one that is given along with the function. For example, the function given by

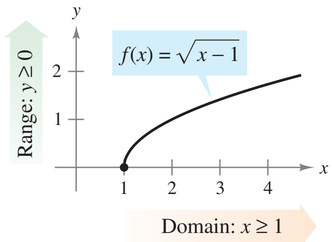
$$f(x) = \frac{1}{x^2 - 4}, \qquad 4 \le x \le 5$$

has an explicitly defined domain given by $\{x:\ 4 \le x \le 5\}$. On the other hand, the function given by

$$g(x) = \frac{1}{x^2 - 4}$$

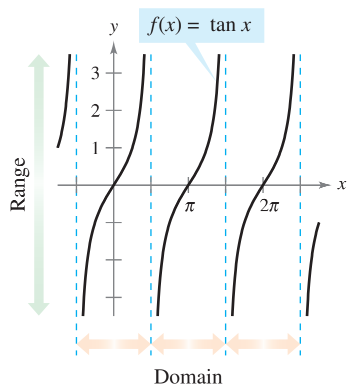
has an implied domain that is the set $\{x:\ x \ne \pm 2\}$.

Video



(a) The domain of f is $[1, \infty)$ and the range is $[0, \infty)$.

Editable Graph



(b) The domain of f is all x-values such that $x \ne \frac{\pi}{2} + n\pi$ and the range is $(-\infty, \infty)$.

Editable Graph

Figure P.23

EXAMPLE 2 Finding the Domain and Range of a Function

a. The domain of the function

$$f(x) = \sqrt{x - 1}$$

is the set of all x-values for which $x - 1 \ge 0$, which is the interval $[1, \infty)$. To find the range observe that $f(x) = \sqrt{x-1}$ is never negative. So, the range is the interval $[0, \infty)$, as indicated in Figure P.23(a).

b. The domain of the tangent function, as shown in Figure P.23(b),

$$f(x) = \tan x$$

is the set of all x-values such that

$$x \ne \frac{\pi}{2} + n\pi, \qquad n \text{ is an integer.} \qquad \text{Domain of tangent function}$$

The range of this function is the set of all real numbers. For a review of the characteristics of this and other trigonometric functions, see Appendix D.

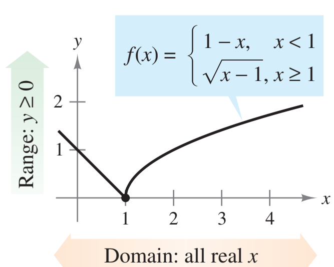
Try It   Exploration A

EXAMPLE 3 A Function Defined by More than One Equation

Determine the domain and range of the function.

$$f(x) = \begin{cases} 1 - x, & \text{if } x < 1 \\ \sqrt{x - 1}, & \text{if } x \ge 1 \end{cases}$$

Solution Because f is defined for $x < 1$ and $x \ge 1$, the domain is the entire set of real numbers. On the portion of the domain for which $x \ge 1$, the function behaves as in Example 2(a). For $x < 1$, the values of $1 - x$ are positive. So, the range of the function is the interval $[0, \infty)$. (See Figure P.24.)



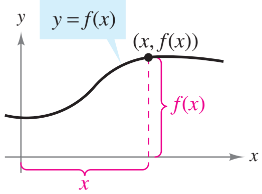
The domain of f is $(-\infty, \infty)$ and the range is $[0, \infty)$.
Figure P.24

Editable Graph

Try It   Exploration A

A function from X to Y is **one-to-one** if to each y-value in the range there corresponds exactly one x-value in the domain. For instance, the function given in Example 2(a) is one-to-one, whereas the functions given in Examples 2(b) and 3 are not one-to-one. A function from X to Y is **onto** if its range consists of all of Y.

The Graph of a Function

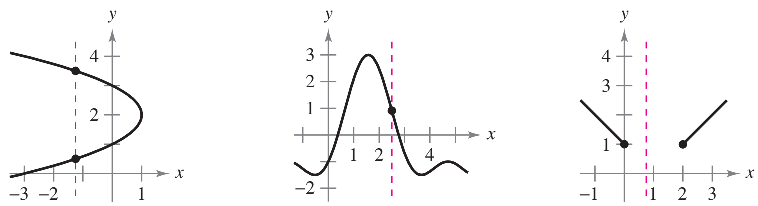


The graph of a function
Figure P.25

The graph of the function $y = f(x)$ consists of all points $(x, f(x))$, where x is in the domain of f. In Figure P.25, note that

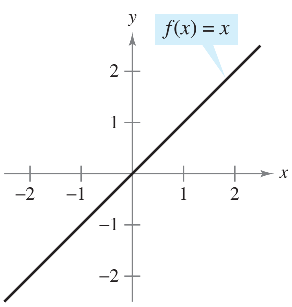
$$x = \text{the directed distance from the } y\text{-axis}$$
$$f(x) = \text{the directed distance from the } x\text{-axis.}$$

A vertical line can intersect the graph of a function of x at most *once*. This observation provides a convenient visual test, called the **Vertical Line Test,** for functions of x. That is, a graph in the coordinate plane is the graph of a function of f if and only if no vertical line intersects the graph at more than one point. For example, in Figure P.26(a), you can see that the graph does not define y as a function of x because a vertical line intersects the graph twice, whereas in Figures P.26(b) and (c), the graphs do define y as a function of x.

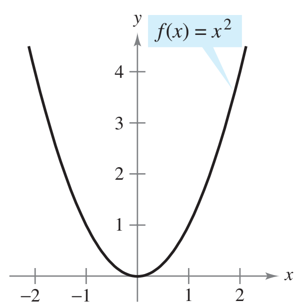


(a) Not a function of x **(b)** A function of x **(c)** A function of x
Figure P.26

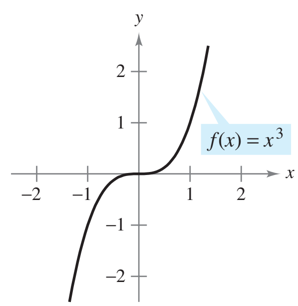
Figure P.27 shows the graphs of eight basic functions. You should be able to recognize these graphs. (Graphs of the other four basic trigonometric functions are shown in Appendix D.)



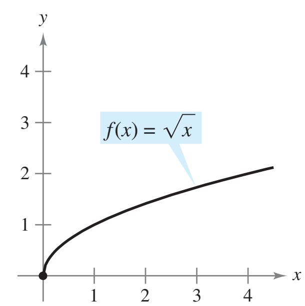
Identity function



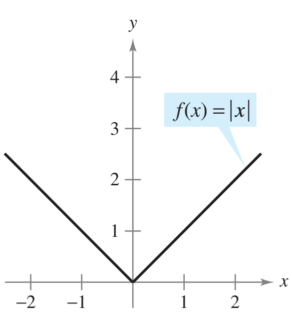
Squaring function



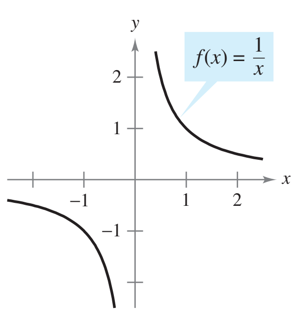
Cubing function



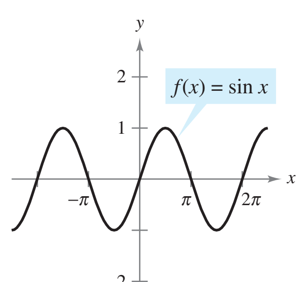
Square root function



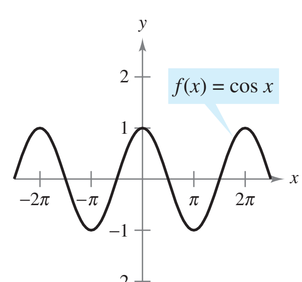
Absolute value function



Rational function



Sine function

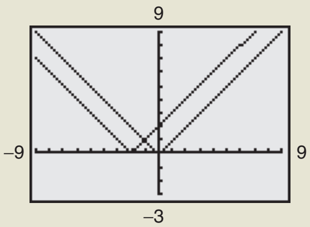


Cosine function

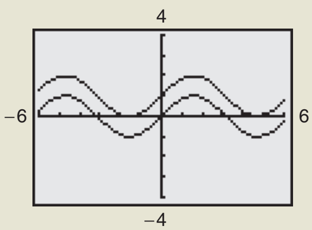
The graphs of eight basic functions
Figure P.27

EXPLORATION

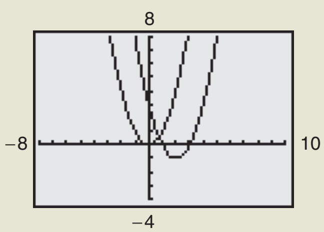
Writing Equations for Functions Each of the graphing utility screens below shows the graph of one of the eight basic functions shown on page 22. Each screen also shows a transformation of the graph. Describe the transformation. Then use your description to write an equation for the transformation.



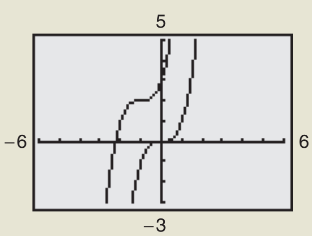
a.



b.



c.



d.

Video

Transformations of Functions

Some families of graphs have the same basic shape. For example, compare the graph of $y = x^2$ with the graphs of the four other quadratic functions shown in Figure P.28.

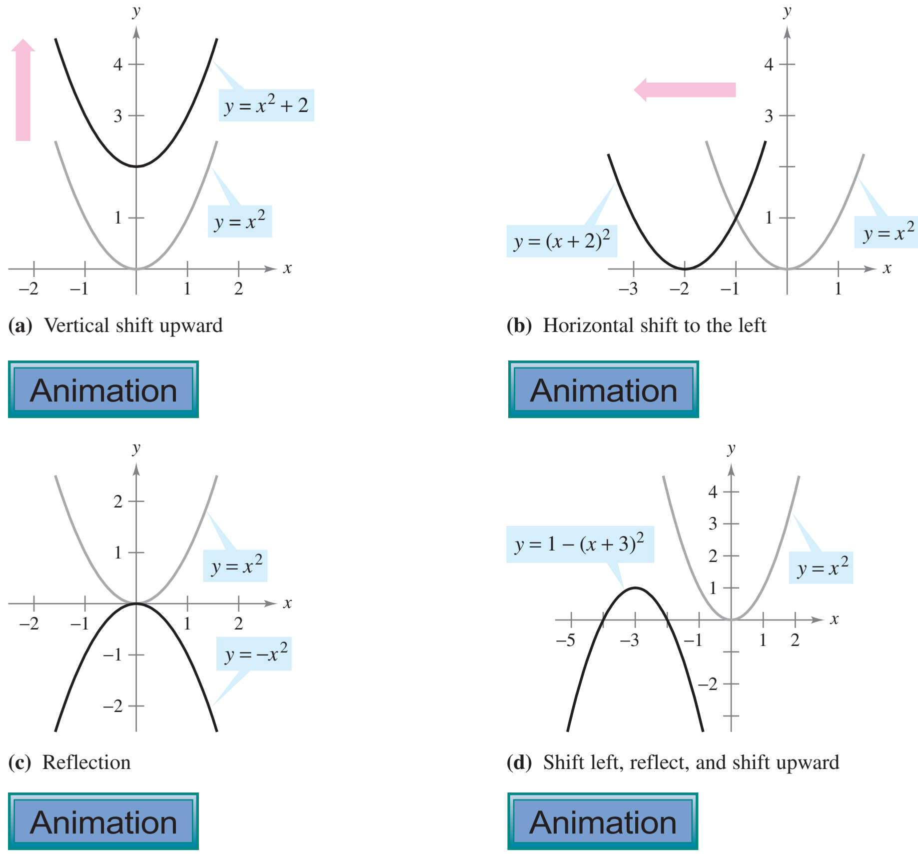


Figure P.28

Each of the graphs in Figure P.28 is a **transformation** of the graph of $y = x^2$. The three basic types of transformations illustrated by these graphs are vertical shifts, horizontal shifts, and reflections. Function notation lends itself well to describing transformations of graphs in the plane. For instance, if $f(x) = x^2$ is considered to be the original function in Figure P.28, the transformations shown can be represented by the following equations.

$y = f(x) + 2$	Vertical shift up 2 units
$y = f(x + 2)$	Horizontal shift to the left 2 units
$y = -f(x)$	Reflection about the x-axis
$y = -f(x + 3) + 1$	Shift left 3 units, reflect about x-axis, and shift up 1 unit

Basic Types of Transformations ($c > 0$)

Original graph:	$y = f(x)$
Horizontal shift c units to the **right**:	$y = f(x - c)$
Horizontal shift c units to the **left**:	$y = f(x + c)$
Vertical shift c units **downward**:	$y = f(x) - c$
Vertical shift c units **upward**:	$y = f(x) + c$
Reflection (about the x-axis):	$y = -f(x)$
Reflection (about the y-axis):	$y = f(-x)$
Reflection (about the origin):	$y = -f(-x)$

Classifications and Combinations of Functions

LEONHARD EULER (1707–1783)

In addition to making major contributions to almost every branch of mathematics, Euler was one of the first to apply calculus to real-life problems in physics. His extensive published writings include such topics as shipbuilding, acoustics, optics, astronomy, mechanics, and magnetism.

MathBio

The modern notion of a function is derived from the efforts of many seventeenth- and eighteenth-century mathematicians. Of particular note was Leonhard Euler, to whom we are indebted for the function notation $y = f(x)$. By the end of the eighteenth century, mathematicians and scientists had concluded that many real-world phenomena could be represented by mathematical models taken from a collection of functions called **elementary functions.** Elementary functions fall into three categories.

1. Algebraic functions (polynomial, radical, rational)
2. Trigonometric functions (sine, cosine, tangent, and so on)
3. Exponential and logarithmic functions

You can review the trigonometric functions in Appendix D. The other nonalgebraic functions, such as the inverse trigonometric functions and the exponential and logarithmic functions, are introduced in Chapter 5.

The most common type of algebraic function is a **polynomial function**

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdot \cdot \cdot + a_2 x^2 + a_1 x + a_0, \qquad a_n \neq 0$$

where the positive integer n is the **degree** of the polynomial function. The constants a_i are **coefficients,** with a_n the **leading coefficient** and a_0 the **constant term** of the polynomial function. It is common practice to use subscript notation for coefficients of general polynomial functions, but for polynomial functions of low degree, the following simpler forms are often used.

Zeroth degree:	$f(x) = a$	Constant function
First degree:	$f(x) = ax + b$	Linear function
Second degree:	$f(x) = ax^2 + bx + c$	Quadratic function
Third degree:	$f(x) = ax^3 + bx^2 + cx + d$	Cubic function

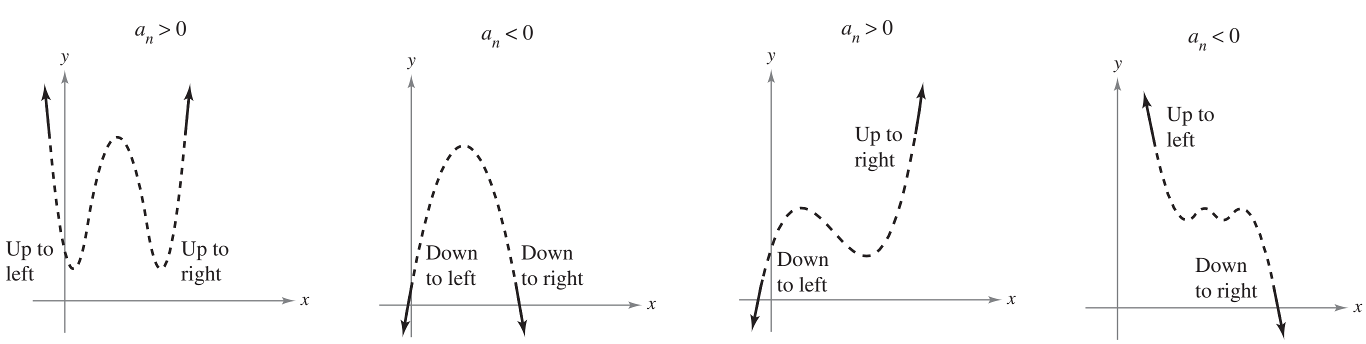
FOR FURTHER INFORMATION For more on the history of the concept of a function, see the article "Evolution of the Function Concept: A Brief Survey" by Israel Kleiner in *The College Mathematics Journal*.

MathArticle

Although the graph of a nonconstant polynomial function can have several turns, eventually the graph will rise or fall without bound as x moves to the right or left. Whether the graph of

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdot \cdot \cdot + a_2 x^2 + a_1 x + a_0$$

eventually rises or falls can be determined by the function's degree (odd or even) and by the leading coefficient a_n, as indicated in Figure P.29. Note that the dashed portions of the graphs indicate that the **Leading Coefficient Test** determines *only* the right and left behavior of the graph.



Graphs of polynomial functions of even degree

Graphs of polynomial functions of odd degree

The Leading Coefficient Test for polynomial functions
Figure P.29

Just as a rational number can be written as the quotient of two integers, a **rational function** can be written as the quotient of two polynomials. Specifically, a function f is rational if it has the form

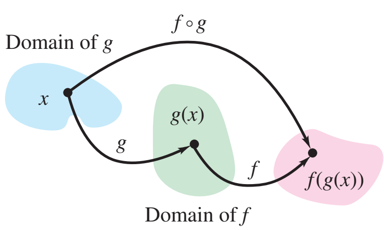
$$f(x) = \frac{p(x)}{q(x)}, \quad q(x) \neq 0$$

where $p(x)$ and $q(x)$ are polynomials.

Polynomial functions and rational functions are examples of **algebraic functions.** An algebraic function of x is one that can be expressed as a finite number of sums, differences, multiples, quotients, and radicals involving x^n. For example, $f(x) = \sqrt{x + 1}$ is algebraic. Functions that are not algebraic are **transcendental.** For instance, the trigonometric functions are transcendental.

Two functions can be combined in various ways to create new functions. For example, given $f(x) = 2x - 3$ and $g(x) = x^2 + 1$, you can form the functions shown.

$$(f + g)(x) = f(x) + g(x) = (2x - 3) + (x^2 + 1) \qquad \text{Sum}$$
$$(f - g)(x) = f(x) - g(x) = (2x - 3) - (x^2 + 1) \qquad \text{Difference}$$
$$(fg)(x) = f(x)g(x) = (2x - 3)(x^2 + 1) \qquad \text{Product}$$
$$(f/g)(x) = \frac{f(x)}{g(x)} = \frac{2x - 3}{x^2 + 1} \qquad \text{Quotient}$$



The domain of the composite function $f \circ g$
Figure P.30

You can combine two functions in yet another way, called **composition.** The resulting function is called a **composite function.**

Definition of Composite Function

Let f and g be functions. The function given by $(f \circ g)(x) = f(g(x))$ is called the **composite** of f with g. The domain of $f \circ g$ is the set of all x in the domain of g such that $g(x)$ is in the domain of f (see Figure P.30).

The composite of f with g may not be equal to the composite of g with f.

EXAMPLE 4 Finding Composite Functions

Given $f(x) = 2x - 3$ and $g(x) = \cos x$, find each composite function.

a. $f \circ g$ **b.** $g \circ f$

Solution

a.
$$\begin{aligned}(f \circ g)(x) &= f(g(x)) && \text{Definition of } f \circ g\\ &= f(\cos x) && \text{Substitute } \cos x \text{ for } g(x).\\ &= 2(\cos x) - 3 && \text{Definition of } f(x)\\ &= 2\cos x - 3 && \text{Simplify.}\end{aligned}$$

b.
$$\begin{aligned}(g \circ f)(x) &= g(f(x)) && \text{Definition of } g \circ f\\ &= g(2x - 3) && \text{Substitute } 2x - 3 \text{ for } f(x).\\ &= \cos(2x - 3) && \text{Definition of } g(x)\end{aligned}$$

Note that $(f \circ g)(x) \neq (g \circ f)(x)$.

Try It | Exploration A | Exploration B

Exploration C | Exploration D | Exploration E | Open Exploration

EXPLORATION

Use a graphing utility to graph each function. Determine whether the function is *even*, *odd*, or *neither*.

$f(x) = x^2 - x^4$

$g(x) = 2x^3 + 1$

$h(x) = x^5 - 2x^3 + x$

$j(x) = 2 - x^6 - x^8$

$k(x) = x^5 - 2x^4 + x - 2$

$p(x) = x^9 + 3x^5 - x^3 + x$

Describe a way to identify a function as odd or even by inspecting the equation.

In Section P.1, an x-intercept of a graph was defined to be a point $(a, 0)$ at which the graph crosses the x-axis. If the graph represents a function f, the number a is a **zero** of f. In other words, *the zeros of a function f are the solutions of the equation $f(x) = 0$.* For example, the function $f(x) = x - 4$ has a zero at $x = 4$ because $f(4) = 0$.

In Section P.1 you also studied different types of symmetry. In the terminology of functions, a function is **even** if its graph is symmetric with respect to the y-axis, and is **odd** if its graph is symmetric with respect to the origin. The symmetry tests in Section P.1 yield the following test for even and odd functions.

Test for Even and Odd Functions

The function $y = f(x)$ is **even** if $f(-x) = f(x)$.

The function $y = f(x)$ is **odd** if $f(-x) = -f(x)$.

NOTE Except for the constant function $f(x) = 0$, the graph of a function of x cannot have symmetry with respect to the x-axis because it then would fail the Vertical Line Test for the graph of the function.

EXAMPLE 5 Even and Odd Functions and Zeros of Functions

Determine whether each function is even, odd, or neither. Then find the zeros of the function.

a. $f(x) = x^3 - x$ **b.** $g(x) = 1 + \cos x$

Solution

a. This function is odd because

$$f(-x) = (-x)^3 - (-x) = -x^3 + x = -(x^3 - x) = -f(x).$$

The zeros of f are found as shown.

$$x^3 - x = 0 \qquad \text{Let } f(x) = 0.$$

$$x(x^2 - 1) = x(x - 1)(x + 1) = 0 \qquad \text{Factor.}$$

$$x = 0, 1, -1 \qquad \text{Zeros of } f$$

See Figure P.31(a).

b. This function is even because

$$g(-x) = 1 + \cos(-x) = 1 + \cos x = g(x). \qquad \cos(-x) = \cos(x)$$

The zeros of g are found as shown.

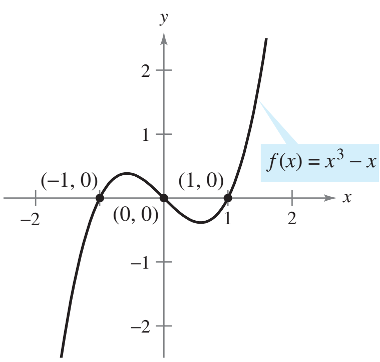
$$1 + \cos x = 0 \qquad \text{Let } g(x) = 0.$$

$$\cos x = -1 \qquad \text{Subtract 1 from each side.}$$

$$x = (2n + 1)\pi, \ n \text{ is an integer.} \qquad \text{Zeros of } g$$

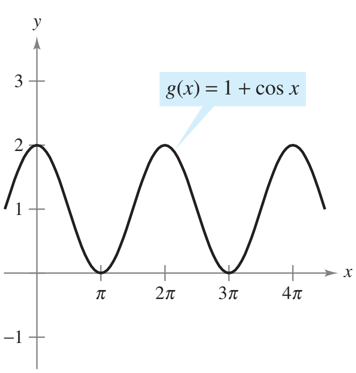
See Figure P.31(b).

Try It | Exploration A | Exploration B | Exploration C



(a) Odd function

Editable Graph



(b) Even function

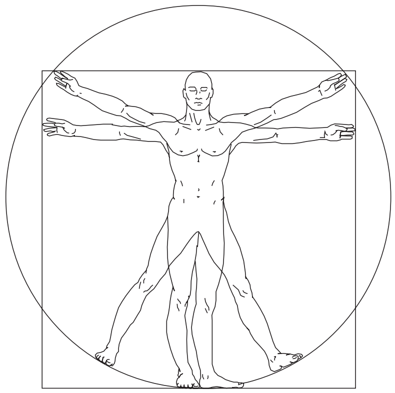
Editable Graph

Figure P.31

NOTE Each of the functions in Example 5 is either even or odd. However, some functions, such as $f(x) = x^2 + x + 1$, are neither even nor odd.

Section P.4 Fitting Models to Data

- **Fit a linear model to a real-life data set.**
- **Fit a quadratic model to a real-life data set.**
- **Fit a trigonometric model to a real-life data set.**



A computer graphics drawing based on the pen and ink drawing of Leonardo da Vinci's famous study of human proportions, called *Vitruvian Man*

Fitting a Linear Model to Data

A basic premise of science is that much of the physical world can be described mathematically and that many physical phenomena are predictable. This scientific outlook was part of the scientific revolution that took place in Europe during the late 1500s. Two early publications connected with this revolution were *On the Revolutions of the Heavenly Spheres* by the Polish astronomer Nicolaus Copernicus and *On the Structure of the Human Body* by the Belgian anatomist Andreas Vesalius. Each of these books was published in 1543 and each broke with prior tradition by suggesting the use of a scientific method rather than unquestioned reliance on authority.

One basic technique of modern science is gathering data and then describing the data with a mathematical model. For instance, the data given in Example 1 are inspired by Leonardo da Vinci's famous drawing that indicates that a person's height and arm span are equal.

EXAMPLE 1 Fitting a Linear Model to Data

A class of 28 people collected the following data, which represent their heights x and arm spans y (rounded to the nearest inch).

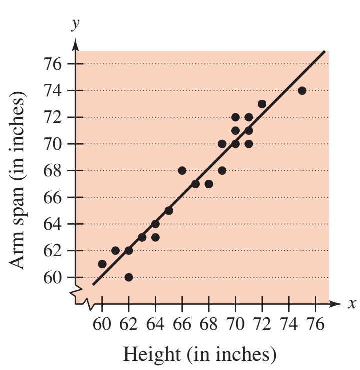
(60, 61), (65, 65), (68, 67), (72, 73), (61, 62), (63, 63), (70, 71),
(75, 74), (71, 72), (62, 60), (65, 65), (66, 68), (62, 62), (72, 73),
(70, 70), (69, 68), (69, 70), (60, 61), (63, 63), (64, 64), (71, 71),
(68, 67), (69, 70), (70, 72), (65, 65), (64, 63), (71, 70), (67, 67)

Find a linear model to represent these data.

Solution There are different ways to model these data with an equation. The simplest would be to observe that x and y are about the same and list the model as simply $y = x$. A more careful analysis would be to use a procedure from statistics called linear regression. (You will study this procedure in Section 13.9.) The least squares regression line for these data is

$$y = 1.006x - 0.23. \qquad \text{Least squares regression line}$$

The graph of the model and the data are shown in Figure P.32. From this model, you can see that a person's arm span tends to be about the same as his or her height.



Linear model and data
Figure P.32

Try It | Exploration A | Open Exploration

Video

TECHNOLOGY Many scientific and graphing calculators have built-in least squares regression programs. Typically, you enter the data into the calculator and then run the linear regression program. The program usually displays the slope and y-intercept of the best-fitting line and the *correlation coefficient* r. The correlation coefficient gives a measure of how well the model fits the data. The closer $|r|$ is to 1, the better the model fits the data. For instance, the correlation coefficient for the model in Example 1 is $r \approx 0.97$, which indicates that the model is a good fit for the data. If the r-value is positive, the variables have a positive correlation, as in Example 1. If the r-value is negative, the variables have a negative correlation.

Fitting a Quadratic Model to Data

A function that gives the height s of a falling object in terms of the time t is called a position function. If air resistance is not considered, the position of a falling object can be modeled by

$$s(t) = \frac{1}{2}gt^2 + v_0t + s_0$$

where g is the acceleration due to gravity, v_0 is the initial velocity, and s_0 is the initial height. The value of g depends on where the object is dropped. On earth, g is approximately -32 feet per second per second, or -9.8 meters per second per second.

To discover the value of g experimentally, you could record the heights of a falling object at several increments, as shown in Example 2.

EXAMPLE 2 Fitting a Quadratic Model to Data

A basketball is dropped from a height of about $5\frac{1}{4}$ feet. The height of the basketball is recorded 23 times at intervals of about 0.02 second.* The results are shown in the table.

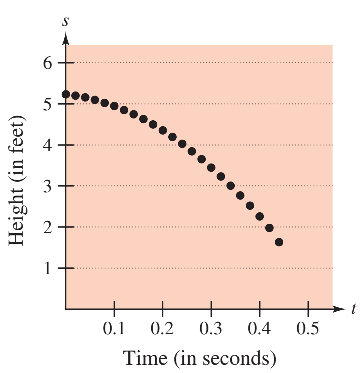
Time	0.0	0.02	0.04	0.06	0.08	0.099996
Height	5.23594	5.20353	5.16031	5.0991	5.02707	4.95146

Time	0.119996	0.139992	0.159988	0.179988	0.199984	0.219984
Height	4.85062	4.74979	4.63096	4.50132	4.35728	4.19523

Time	0.23998	0.25993	0.27998	0.299976	0.319972	0.339961
Height	4.02958	3.84593	3.65507	3.44981	3.23375	3.01048

Time	0.359961	0.379951	0.399941	0.419941	0.439941
Height	2.76921	2.52074	2.25786	1.98058	1.63488

Find a model to fit these data. Then use the model to predict the time when the basketball will hit the ground.



Scatter plot of data
Figure P.33

Solution Begin by drawing a scatter plot of the data, as shown in Figure P.33. From the scatter plot, you can see that the data do not appear to be linear. It does appear, however, that they might be quadratic. To check this, enter the data into a calculator or computer that has a quadratic regression program. You should obtain the model

$$s = -15.45t^2 - 1.30t + 5.234. \quad \text{Least squares regression quadratic}$$

Using this model, you can predict the time when the basketball hits the ground by substituting 0 for s and solving the resulting equation for t.

$$0 = -15.45t^2 - 1.30t + 5.234 \quad \text{Let } s = 0.$$

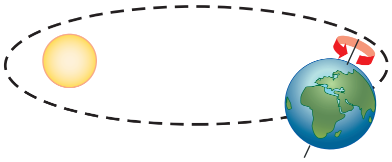
$$t = \frac{1.30 \pm \sqrt{(-1.30)^2 - 4(-15.45)(5.234)}}{2(-15.45)} \quad \text{Quadratic Formula}$$

$$t \approx 0.54 \quad \text{Choose positive solution.}$$

The solution is about 0.54 second. In other words, the basketball will continue to fall for about 0.1 second more before hitting the ground.

Try It | Exploration A

* *Data were collected with a Texas Instruments CBL (Calculator-Based Laboratory) System.*



The plane of Earth's orbit about the sun and its axis of rotation are not perpendicular. Instead, Earth's axis is tilted with respect to its orbit. The result is that the amount of daylight received by locations on Earth varies with the time of year. That is, it varies with the position of Earth in its orbit.

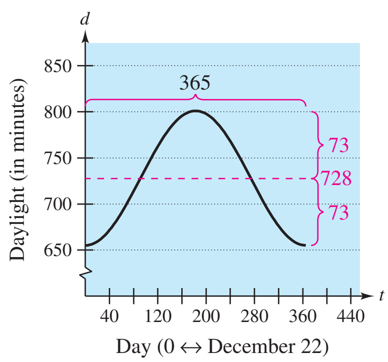
Fitting a Trigonometric Model to Data

What is mathematical modeling? This is one of the questions that is asked in the book *Guide to Mathematical Modelling*. Here is part of the answer.*

1. Mathematical modeling consists of applying your mathematical skills to obtain useful answers to real problems.
2. Learning to apply mathematical skills is very different from learning mathematics itself.
3. Models are used in a very wide range of applications, some of which do not appear initially to be mathematical in nature.
4. Models often allow quick and cheap evaluation of alternatives, leading to optimal solutions that are not otherwise obvious.
5. There are no precise rules in mathematical modeling and no "correct" answers.
6. Modeling can be learned only by *doing*.

EXAMPLE 3 Fitting a Trigonometric Model to Data

The number of hours of daylight on Earth depends on the latitude and the time of year. Here are the numbers of minutes of daylight at a location of 20° N latitude on the longest and shortest days of the year: June 21, 801 minutes; December 22, 655 minutes. Use these data to write a model for the amount of daylight d (in minutes) on each day of the year at a location of 20° N latitude. How could you check the accuracy of your model?



Graph of model
Figure P.34

Solution Here is one way to create a model. You can hypothesize that the model is a sine function whose period is 365 days. Using the given data, you can conclude that the amplitude of the graph is $(801 - 655)/2$, or 73. So, one possible model is

$$d = 728 - 73\sin\left(\frac{2\pi t}{365} + \frac{\pi}{2}\right).$$

In this model, t represents the number of each day of the year, with December 22 represented by $t = 0$. A graph of this model is shown in Figure P.34. To check the accuracy of this model, we used a weather almanac to find the numbers of minutes of daylight on different days of the year at the location of 20° N latitude.

Date	*Value of t*	*Actual Daylight*	*Daylight Given by Model*
Dec 22	0	655 min	655 min
Jan 1	10	657 min	656 min
Feb 1	41	676 min	672 min
Mar 1	69	705 min	701 min
Apr 1	100	740 min	739 min
May 1	130	772 min	773 min
Jun 1	161	796 min	796 min
Jun 21	181	801 min	801 min
Jul 1	191	799 min	800 min
Aug 1	222	782 min	785 min
Sep 1	253	752 min	754 min
Oct 1	283	718 min	716 min
Nov 1	314	685 min	681 min
Dec 1	344	661 min	660 min

You can see that the model is fairly accurate.

NOTE For more review of trigonometric functions, see Appendix D.

Try It **Exploration A**

* *Text from Dilwyn Edwards and Mike Hamson,* Guide to Mathematical Modelling *(Boca Raton: CRC Press, 1990). Used by permission of the authors.*

Section 1.1 A Preview of Calculus

- **Understand what calculus is and how it compares with precalculus.**
- **Understand that the tangent line problem is basic to calculus.**
- **Understand that the area problem is also basic to calculus.**

STUDY TIP As you progress through this course, remember that learning calculus is just one of your goals. Your most important goal is to learn how to use calculus to model and solve real-life problems. Here are a few problem-solving strategies that may help you.

- Be sure you understand the question. What is given? What are you asked to find?
- Outline a plan. There are many approaches you could use: look for a pattern, solve a simpler problem, work backwards, draw a diagram, use technology, or any of many other approaches.
- Complete your plan. Be sure to answer the question. Verbalize your answer. For example, rather than writing the answer as $x = 4.6$, it would be better to write the answer as "The area of the region is 4.6 square meters."
- Look back at your work. Does your answer make sense? Is there a way you can check the reasonableness of your answer?

What Is Calculus?

Calculus is the mathematics of change—velocities and accelerations. Calculus is also the mathematics of tangent lines, slopes, areas, volumes, arc lengths, centroids, curvatures, and a variety of other concepts that have enabled scientists, engineers, and economists to model real-life situations.

Although precalculus mathematics also deals with velocities, accelerations, tangent lines, slopes, and so on, there is a fundamental difference between precalculus mathematics and calculus. Precalculus mathematics is more static, whereas calculus is more dynamic. Here are some examples.

- An object traveling at a constant velocity can be analyzed with precalculus mathematics. To analyze the velocity of an accelerating object, you need calculus.
- The slope of a line can be analyzed with precalculus mathematics. To analyze the slope of a curve, you need calculus.
- A tangent line to a circle can be analyzed with precalculus mathematics. To analyze a tangent line to a general graph, you need calculus.
- The area of a rectangle can be analyzed with precalculus mathematics. To analyze the area under a general curve, you need calculus.

Each of these situations involves the same general strategy—the reformulation of precalculus mathematics through the use of a limit process. So, one way to answer the question "What is calculus?" is to say that calculus is a "limit machine" that involves three stages. The first stage is precalculus mathematics, such as the slope of a line or the area of a rectangle. The second stage is the limit process, and the third stage is a new calculus formulation, such as a derivative or integral.

Precalculus mathematics ⇨ Limit process ⇨ Calculus

Some students try to learn calculus as if it were simply a collection of new formulas. This is unfortunate. If you reduce calculus to the memorization of differentiation and integration formulas, you will miss a great deal of understanding, self-confidence, and satisfaction.

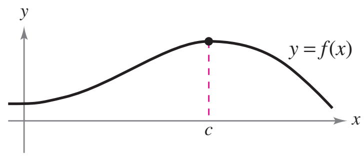
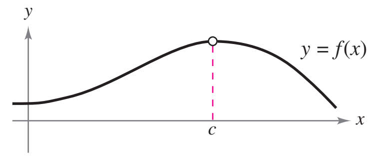
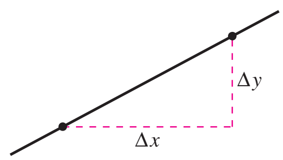
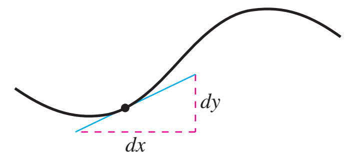
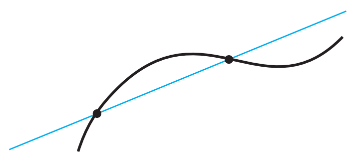
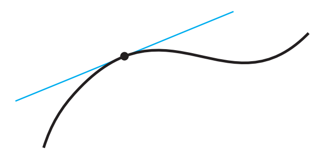
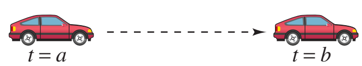
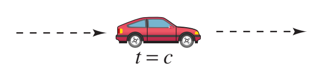
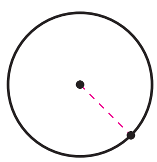
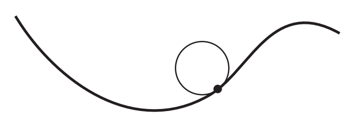
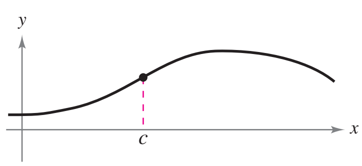
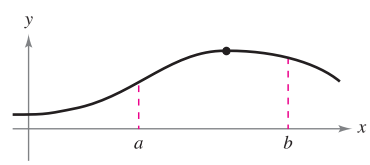
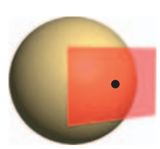
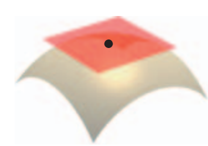
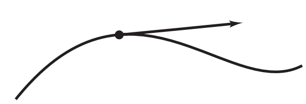
On the following two pages some familiar precalculus concepts coupled with their calculus counterparts are listed. Throughout the text, your goal should be to learn how precalculus formulas and techniques are used as building blocks to produce the more general calculus formulas and techniques. Don't worry if you are unfamiliar with some of the "old formulas" listed on the following two pages—you will be reviewing all of them.

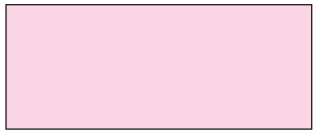
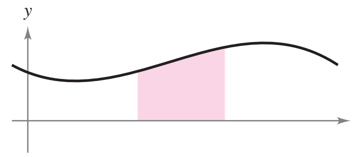
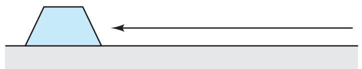
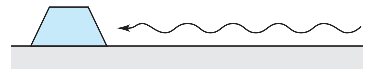
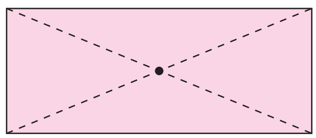
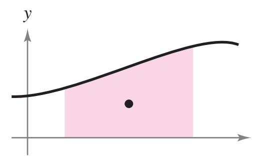
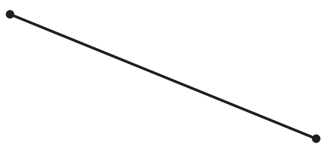
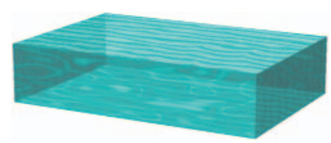
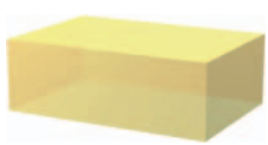
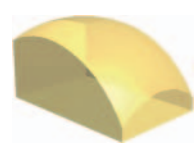
As you proceed through this text, come back to this discussion repeatedly. Try to keep track of where you are relative to the three stages involved in the study of calculus. For example, the first three chapters break down as shown.

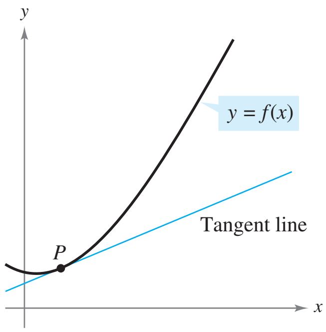
Chapter P: Preparation for Calculus	Precalculus
Chapter 1: Limits and Their Properties	Limit process
Chapter 2: Differentiation	Calculus

GRACE CHISHOLM YOUNG (1868–1944)

Grace Chisholm Young received her degree in mathematics from Girton College in Cambridge, England. Her early work was published under the name of William Young, her husband. Between 1914 and 1916, Grace Young published work on the foundations of calculus that won her the Gamble Prize from Girton College.

Without Calculus		With Differential Calculus	
Value of $f(x)$ when $x = c$		Limit of $f(x)$ as x approaches c	
Slope of a line		Slope of a curve	
Secant line to a curve		Tangent line to a curve	
Average rate of change between $t = a$ and $t = b$		Instantaneous rate of change at $t = c$	
Curvature of a circle		Curvature of a curve	
Height of a curve when $x = c$		Maximum height of a curve on an interval	
Tangent plane to a sphere		Tangent plane to a surface	
Direction of motion along a line		Direction of motion along a curve	

Without Calculus		**With Integral Calculus**	
Area of a rectangle		Area under a curve	
Work done by a constant force		Work done by a variable force	
Center of a rectangle		Centroid of a region	
Length of a line segment		Length of an arc	
Surface area of a cylinder		Surface area of a solid of revolution	
Mass of a solid of constant density		Mass of a solid of variable density	
Volume of a rectangular solid		Volume of a region under a surface	
Sum of a finite number of terms	$a_1 + a_2 + \cdots + a_n = S$	Sum of an infinite number of terms	$a_1 + a_2 + a_3 + \cdots = S$



The tangent line to the graph of f at P
Figure 1.1

Video

The Tangent Line Problem

The notion of a limit is fundamental to the study of calculus. The following brief descriptions of two classic problems in calculus—*the tangent line problem and the area problem*—should give you some idea of the way limits are used in calculus.

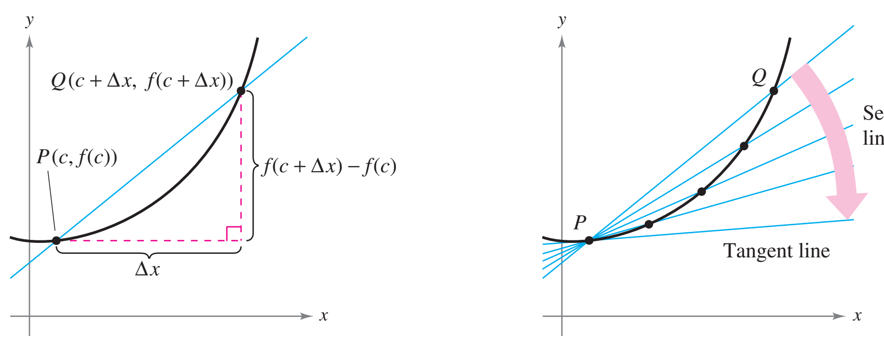
In the tangent line problem, you are given a function f and a point P on its graph and are asked to find an equation of the tangent line to the graph at point P, as shown in Figure 1.1.

Except for cases involving a vertical tangent line, the problem of finding the **tangent line** at a point P is equivalent to finding the *slope* of the tangent line at P. You can approximate this slope by using a line through the point of tangency and a second point on the curve, as shown in Figure 1.2(a). Such a line is called a **secant line.** If $P(c, f(c))$ is the point of tangency and

$$Q(c + \Delta x, f(c + \Delta x))$$

is a second point on the graph of f, the slope of the secant line through these two points is given by

$$m_{sec} = \frac{f(c + \Delta x) - f(c)}{c + \Delta x - c} = \frac{f(c + \Delta x) - f(c)}{\Delta x}.$$



(a) The secant line through $(c, f(c))$ and $(c + \Delta x, f(c + \Delta x))$

(b) As Q approaches P, the secant lines approach the tangent line.

Figure 1.2

Animation

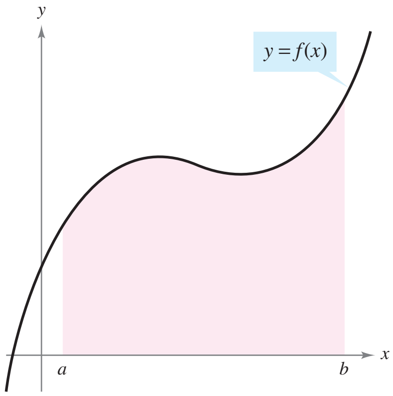
As point Q approaches point P, the slope of the secant line approaches the slope of the tangent line, as shown in Figure 1.2(b). When such a "limiting position" exists, the slope of the tangent line is said to be the **limit** of the slope of the secant line. (Much more will be said about this important problem in Chapter 2.)

EXPLORATION

The following points lie on the graph of $f(x) = x^2$.

$$Q_1(1.5, f(1.5)), \quad Q_2(1.1, f(1.1)), \quad Q_3(1.01, f(1.01)),$$
$$Q_4(1.001, f(1.001)), \quad Q_5(1.0001, f(1.0001))$$

Each successive point gets closer to the point $P(1, 1)$. Find the slope of the secant line through Q_1 and P, Q_2 and P, and so on. Graph these secant lines on a graphing utility. Then use your results to estimate the slope of the tangent line to the graph of f at the point P.



Area under a curve
Figure 1.3

Video

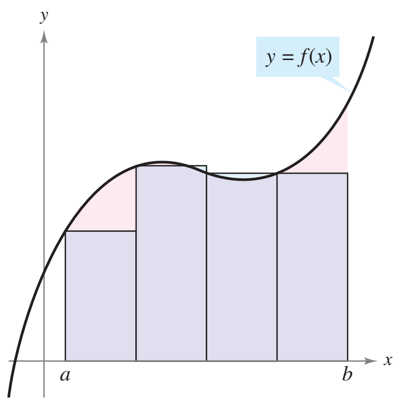
Historical Note

In one of the most astounding events ever to occur in mathematics, it was discovered that the tangent line problem and the area problem are closely related. This discovery led to the birth of calculus. You will learn about the relationship between these two problems when you study the Fundamental Theorem of Calculus in Chapter 4.

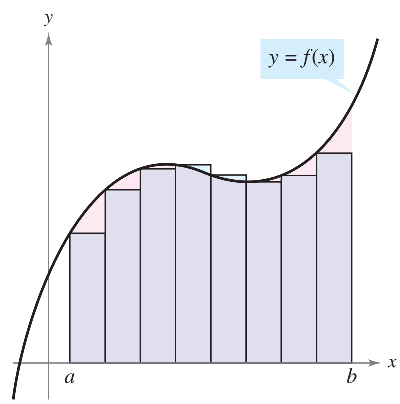
The Area Problem

In the tangent line problem, you saw how the limit process can be applied to the slope of a line to find the slope of a general curve. A second classic problem in calculus is finding the area of a plane region that is bounded by the graphs of functions. This problem can also be solved with a limit process. In this case, the limit process is applied to the area of a rectangle to find the area of a general region.

As a simple example, consider the region bounded by the graph of the function $y = f(x)$, the x-axis, and the vertical lines $x = a$ and $x = b$, as shown in Figure 1.3. You can approximate the area of the region with several rectangular regions, as shown in Figure 1.4. As you increase the number of rectangles, the approximation tends to become better and better because the amount of area missed by the rectangles decreases. Your goal is to determine the limit of the sum of the areas of the rectangles as the number of rectangles increases without bound.



Approximation using four rectangles
Figure 1.4

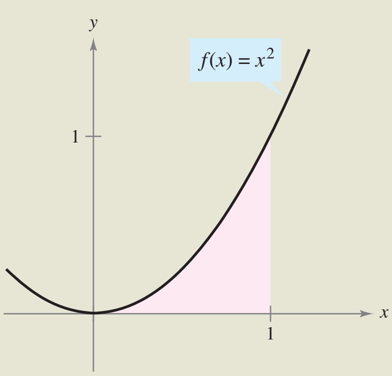


Approximation using eight rectangles

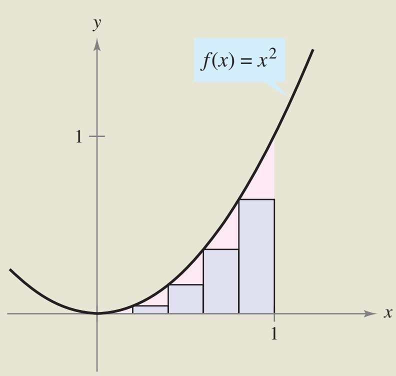
Animation

EXPLORATION

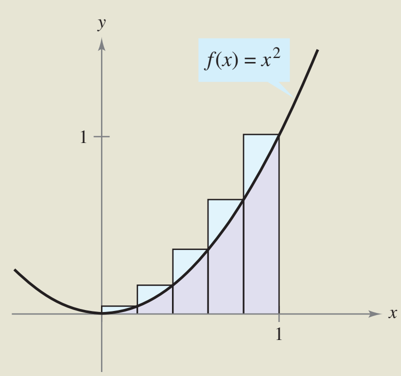
Consider the region bounded by the graphs of $f(x) = x^2$, $y = 0$, and $x = 1$, as shown in part (a) of the figure. The area of the region can be approximated by two sets of rectangles—one set inscribed within the region and the other set circumscribed over the region, as shown in parts (b) and (c). Find the sum of the areas of each set of rectangles. Then use your results to approximate the area of the region.



(a) Bounded region



(b) Inscribed rectangles



(c) Circumscribed rectangles

Section 1.2

Finding Limits Graphically and Numerically

- **Estimate a limit using a numerical or graphical approach.**
- **Learn different ways that a limit can fail to exist.**
- **Study and use a formal definition of limit.**

Video

An Introduction to Limits

Suppose you are asked to sketch the graph of the function f given by

$$f(x) = \frac{x^3 - 1}{x - 1}, \qquad x \neq 1.$$

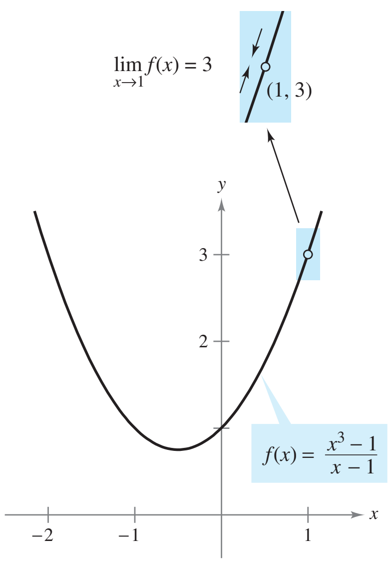
For all values other than $x = 1$, you can use standard curve-sketching techniques. However, at $x = 1$, it is not clear what to expect. To get an idea of the behavior of the graph of f near $x = 1$, you can use two sets of x-values—one set that approaches 1 from the left and one set that approaches 1 from the right, as shown in the table.

x approaches 1 from the left. → ← x approaches 1 from the right.

x	0.75	0.9	0.99	0.999	1	1.001	1.01	1.1	1.25
$f(x)$	2.313	2.710	2.970	2.997	?	3.003	3.030	3.310	3.813

$f(x)$ approaches 3. → ← $f(x)$ approaches 3.

Animation



The limit of $f(x)$ as x approaches 1 is 3.
Figure 1.5

Animation

The graph of f is a parabola that has a gap at the point $(1, 3)$, as shown in Figure 1.5. Although x cannot equal 1, you can move arbitrarily close to 1, and as a result $f(x)$ moves arbitrarily close to 3. Using limit notation, you can write

$$\lim_{x\to 1} f(x) = 3. \qquad \text{This is read as "the limit of } f(x) \text{ as } x \text{ approaches 1 is 3."}$$

This discussion leads to an informal description of a limit. If $f(x)$ becomes arbitrarily close to a single number L as x approaches c from either side, the **limit** of $f(x)$, as x approaches c, is L. This limit is written as

$$\lim_{x\to c} f(x) = L.$$

EXPLORATION

The discussion above gives an example of how you can estimate a limit *numerically* by constructing a table and *graphically* by drawing a graph. Estimate the following limit numerically by completing the table.

$$\lim_{x\to 2} \frac{x^2 - 3x + 2}{x - 2}$$

x	1.75	1.9	1.99	1.999	2	2.001	2.01	2.1	2.25
$f(x)$	?	?	?	?	?	?	?	?	?

Then use a graphing utility to estimate the limit graphically.

EXAMPLE 1 Estimating a Limit Numerically

Evaluate the function $f(x) = x/\left(\sqrt{x+1} - 1\right)$ at several points near $x = 0$ and use the results to estimate the limit

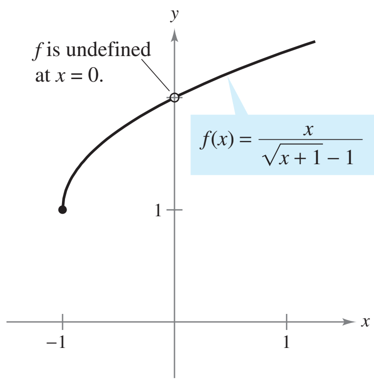
$$\lim_{x \to 0} \frac{x}{\sqrt{x+1} - 1}.$$

Solution The table lists the values of $f(x)$ for several x-values near 0.

x approaches 0 from the left. x approaches 0 from the right.

x	-0.01	-0.001	-0.0001	0	0.0001	0.001	0.01
$f(x)$	1.99499	1.99950	1.99995	?	2.00005	2.00050	2.00499

$f(x)$ approaches 2. $f(x)$ approaches 2.



The limit of $f(x)$ as x approaches 0 is 2.
Figure 1.6

Editable Graph

From the results shown in the table, you can estimate the limit to be 2. This limit is reinforced by the graph of f (see Figure 1.6).

Try It Exploration A Exploration B

In Example 1, note that the function is undefined at $x = 0$ and yet $f(x)$ appears to be approaching a limit as x approaches 0. This often happens, and it is important to realize that *the existence or nonexistence of $f(x)$ at $x = c$ has no bearing on the existence of the limit of $f(x)$ as x approaches c.*

EXAMPLE 2 Finding a Limit

Find the limit of $f(x)$ as x approaches 2 where f is defined as

$$f(x) = \begin{cases} 1, & x \neq 2 \\ 0, & x = 2. \end{cases}$$

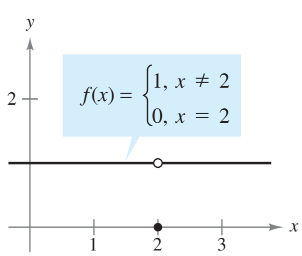
Solution Because $f(x) = 1$ for all x other than $x = 2$, you can conclude that the limit is 1, as shown in Figure 1.7. So, you can write

$$\lim_{x \to 2} f(x) = 1.$$

The fact that $f(2) = 0$ has no bearing on the existence or value of the limit as x approaches 2. For instance, if the function were defined as

$$f(x) = \begin{cases} 1, & x \neq 2 \\ 2, & x = 2 \end{cases}$$

the limit would be the same.



The limit of $f(x)$ as x approaches 2 is 1.
Figure 1.7

Editable Graph

Try It Exploration A

So far in this section, you have been estimating limits numerically and graphically. Each of these approaches produces an estimate of the limit. In Section 1.3, you will study analytic techniques for evaluating limits. Throughout the course, try to develop a habit of using this three-pronged approach to problem solving.

1. Numerical approach — Construct a table of values.
2. Graphical approach — Draw a graph by hand or using technology.
3. Analytic approach — Use algebra or calculus.

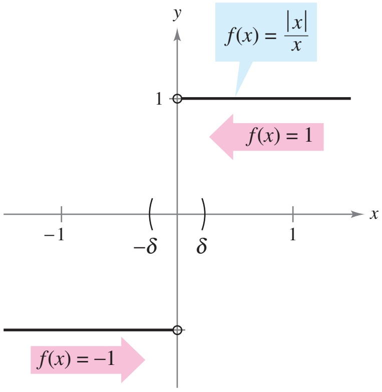
Limits That Fail to Exist

In the next three examples you will examine some limits that fail to exist.

EXAMPLE 3 Behavior That Differs from the Right and Left

Show that the limit does not exist.

$$\lim_{x \to 0} \frac{|x|}{x}$$



$\lim_{x \to 0} f(x)$ does not exist.

Figure 1.8

Editable Graph

Solution Consider the graph of the function $f(x) = |x|/x$. From Figure 1.8, you can see that for positive x-values

$$\frac{|x|}{x} = 1, \qquad x > 0$$

and for negative x-values

$$\frac{|x|}{x} = -1, \qquad x < 0.$$

This means that no matter how close x gets to 0, there will be both positive and negative x-values that yield $f(x) = 1$ and $f(x) = -1$. Specifically, if δ (the lowercase Greek letter *delta*) is a positive number, then for x-values satisfying the inequality $0 < |x| < \delta$, you can classify the values of $|x|/x$ as shown.

$(-\delta, 0)$ $(0, \delta)$

Negative x-values yield $|x|/x = -1$. Positive x-values yield $|x|/x = 1$.

This implies that the limit does not exist.

Try It Exploration A Exploration B

EXAMPLE 4 Unbounded Behavior

Discuss the existence of the limit

$$\lim_{x \to 0} \frac{1}{x^2}.$$

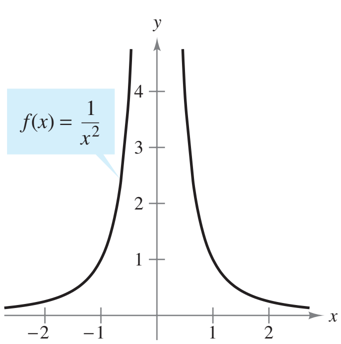
Solution Let $f(x) = 1/x^2$. In Figure 1.9, you can see that as x approaches 0 from either the right or the left, $f(x)$ increases without bound. This means that by choosing x close enough to 0, you can force $f(x)$ to be as large as you want. For instance, $f(x)$ will be larger than 100 if you choose x that is within $\frac{1}{10}$ of 0. That is,

$$0 < |x| < \frac{1}{10} \quad \Rightarrow \quad f(x) = \frac{1}{x^2} > 100.$$

Similarly, you can force $f(x)$ to be larger than 1,000,000, as follows.

$$0 < |x| < \frac{1}{1000} \quad \Rightarrow \quad f(x) = \frac{1}{x^2} > 1{,}000{,}000$$

Because $f(x)$ is not approaching a real number L as x approaches 0, you can conclude that the limit does not exist.



$\lim_{x \to 0} f(x)$ does not exist.

Figure 1.9

Editable Graph

Try It Exploration A Exploration B

EXAMPLE 5 Oscillating Behavior

Discuss the existence of the limit $\lim_{x \to 0} \sin \frac{1}{x}$.

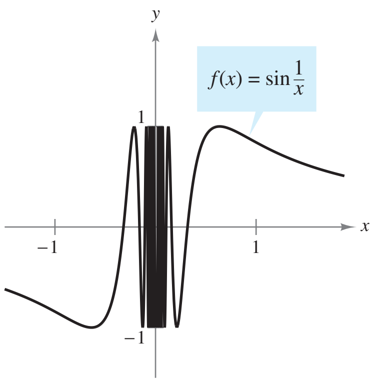
Solution Let $f(x) = \sin(1/x)$. In Figure 1.10, you can see that as x approaches 0, $f(x)$ oscillates between -1 and 1. So, the limit does not exist because no matter how small you choose δ, it is possible to choose x_1 and x_2 within δ units of 0 such that $\sin(1/x_1) = 1$ and $\sin(1/x_2) = -1$, as shown in the table.

x	$2/\pi$	$2/3\pi$	$2/5\pi$	$2/7\pi$	$2/9\pi$	$2/11\pi$	$x \to 0$
sin (1/x)	1	-1	1	-1	1	-1	Limit does not exist.

Try It

Exploration A

Open Exploration



$\lim_{x \to 0} f(x)$ does not exist.
Figure 1.10

Editable Graph

Common Types of Behavior Associated with Nonexistence of a Limit

1. $f(x)$ approaches a different number from the right side of c than it approaches from the left side.
2. $f(x)$ increases or decreases without bound as x approaches c.
3. $f(x)$ oscillates between two fixed values as x approaches c.

There are many other interesting functions that have unusual limit behavior. An often cited one is the *Dirichlet function*

$$f(x) = \begin{cases} 0, & \text{if } x \text{ is rational.} \\ 1, & \text{if } x \text{ is irrational.} \end{cases}$$

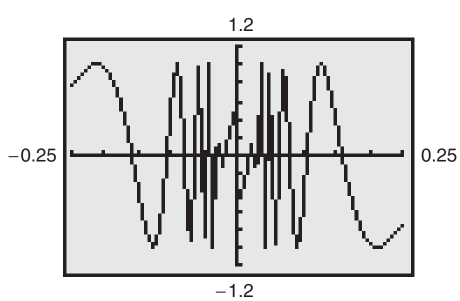
Because this function has *no limit* at any real number c, it is *not continuous* at any real number c. You will study continuity more closely in Section 1.4.

PETER GUSTAV DIRICHLET (1805–1859)

In the early development of calculus, the definition of a function was much more restricted than it is today, and "functions" such as the Dirichlet function would not have been considered. The modern definition of function was given by the German mathematician Peter Gustav Dirichlet.

MathBio

TECHNOLOGY PITFALL When you use a graphing utility to investigate the behavior of a function near the x-value at which you are trying to evaluate a limit, remember that you can't always trust the pictures that graphing utilities draw. If you use a graphing utility to graph the function in Example 5 over an interval containing 0, you will most likely obtain an incorrect graph such as that shown in Figure 1.11. The reason that a graphing utility can't show the correct graph is that the graph has infinitely many oscillations over any interval that contains 0.



Incorrect graph of $f(x) = \sin(1/x)$.
Figure 1.11

A Formal Definition of Limit

Let's take another look at the informal description of a limit. If $f(x)$ becomes arbitrarily close to a single number L as x approaches c from either side, then the limit of $f(x)$ as x approaches c is L, written as

$$\lim_{x \to c} f(x) = L.$$

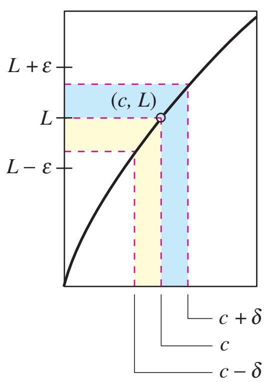
At first glance, this description looks fairly technical. Even so, it is informal because exact meanings have not yet been given to the two phrases

"$f(x)$ becomes arbitrarily close to L"

and

"x approaches c."

The first person to assign mathematically rigorous meanings to these two phrases was Augustin-Louis Cauchy. His **ε-δ definition of limit** is the standard used today.



The ε-δ definition of the limit of $f(x)$ as x approaches c
Figure 1.12

In Figure 1.12, let ε (the lowercase Greek letter *epsilon*) represent a (small) positive number. Then the phrase "$f(x)$ becomes arbitrarily close to L" means that $f(x)$ lies in the interval $(L - \varepsilon, L + \varepsilon)$. Using absolute value, you can write this as

$$|f(x) - L| < \varepsilon.$$

Similarly, the phrase "x approaches c" means that there exists a positive number δ such that x lies in either the interval $(c - \delta, c)$ or the interval $(c, c + \delta)$. This fact can be concisely expressed by the double inequality

$$0 < |x - c| < \delta.$$

The first inequality

$$0 < |x - c| \qquad \text{The distance between } x \text{ and } c \text{ is more than 0.}$$

expresses the fact that $x \neq c$. The second inequality

$$|x - c| < \delta \qquad x \text{ is within } \delta \text{ units of } c.$$

says that x is within a distance δ of c.

Definition of Limit

Let f be a function defined on an open interval containing c (except possibly at c) and let L be a real number. The statement

$$\lim_{x \to c} f(x) = L$$

means that for each $\varepsilon > 0$ there exists a $\delta > 0$ such that if

$$0 < |x - c| < \delta, \quad \text{then} \quad |f(x) - L| < \varepsilon.$$

FOR FURTHER INFORMATION For more on the introduction of rigor to calculus, see "Who Gave You the Epsilon? Cauchy and the Origins of Rigorous Calculus" by Judith V. Grabiner in *The American Mathematical Monthly*.

MathArticle

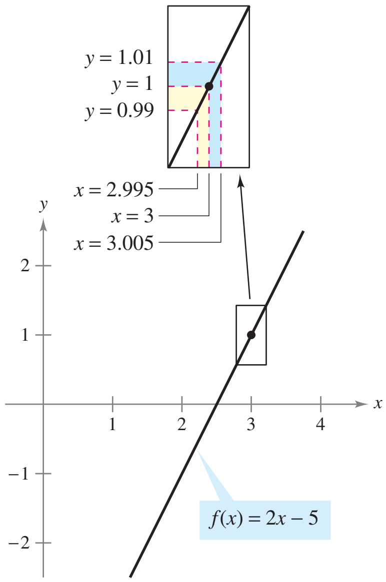
NOTE Throughout this text, the expression

$$\lim_{x \to c} f(x) = L$$

implies two statements—the limit exists *and* the limit is L.

Some functions do not have limits as $x \to c$, but those that do cannot have two different limits as $x \to c$. That is, *if the limit of a function exists, it is unique* (see Exercise 69).

The next three examples should help you develop a better understanding of the ε-δ definition of limit.



The limit of $f(x)$ as x approaches 3 is 1.
Figure 1.13

EXAMPLE 6 Finding a δ for a Given ε

Given the limit

$$\lim_{x\to 3}(2x-5)=1$$

find δ such that $|(2x-5)-1|<0.01$ whenever $0<|x-3|<\delta$.

Solution In this problem, you are working with a given value of ε—namely, $\varepsilon=0.01$. To find an appropriate δ, notice that

$$|(2x-5)-1|=|2x-6|=2|x-3|.$$

Because the inequality $|(2x-5)-1|<0.01$ is equivalent to $2|x-3|<0.01$, you can choose $\delta=\frac{1}{2}(0.01)=0.005$. This choice works because

$$0<|x-3|<0.005$$

implies that

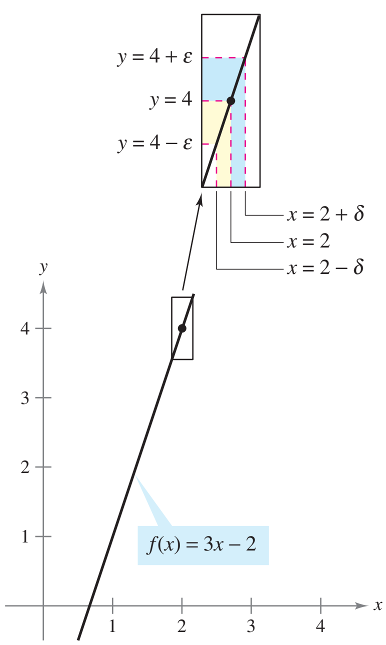
$$|(2x-5)-1|=2|x-3|<2(0.005)=0.01$$

as shown in Figure 1.13.

Try It | Exploration A | Exploration B

NOTE In Example 6, note that 0.005 is the *largest* value of δ that will guarantee $|(2x-5)-1|<0.01$ whenever $0<|x-3|<\delta$. Any *smaller* positive value of δ would also work.

In Example 6, you found a δ-value for a *given* ε. This does not prove the existence of the limit. To do that, you must prove that you can find a δ for any ε, as shown in the next example.



The limit of $f(x)$ as x approaches 2 is 4.
Figure 1.14

EXAMPLE 7 Using the ε-δ Definition of Limit

Use the ε-δ definition of limit to prove that

$$\lim_{x\to 2}(3x-2)=4.$$

Solution You must show that for each $\varepsilon>0$, there exists a $\delta>0$ such that $|(3x-2)-4|<\varepsilon$ whenever $0<|x-2|<\delta$. Because your choice of δ depends on ε, you need to establish a connection between the absolute values $|(3x-2)-4|$ and $|x-2|$.

$$|(3x-2)-4|=|3x-6|=3|x-2|$$

So, for a given $\varepsilon>0$ you can choose $\delta=\varepsilon/3$. This choice works because

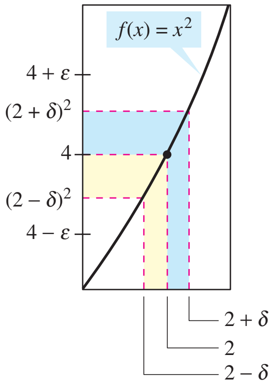
$$0<|x-2|<\delta=\frac{\varepsilon}{3}$$

implies that

$$|(3x-2)-4|=3|x-2|<3\left(\frac{\varepsilon}{3}\right)=\varepsilon$$

as shown in Figure 1.14.

Try It | Exploration A



The limit of $f(x)$ as x approaches 2 is 4.
Figure 1.15

EXAMPLE 8 Using the ε-δ Definition of Limit

Use the ε-δ definition of limit to prove that

$$\lim_{x \to 2} x^2 = 4.$$

Solution You must show that for each $\varepsilon > 0$, there exists a $\delta > 0$ such that

$$|x^2 - 4| < \varepsilon \text{ whenever } 0 < |x - 2| < \delta.$$

To find an appropriate δ, begin by writing $|x^2 - 4| = |x - 2||x + 2|$. For all x in the interval $(1, 3)$, you know that $|x + 2| < 5$. So, letting δ be the minimum of $\varepsilon/5$ and 1, it follows that, whenever $0 < |x - 2| < \delta$, you have

$$|x^2 - 4| = |x - 2||x + 2| < \left(\frac{\varepsilon}{5}\right)(5) = \varepsilon$$

as shown in Figure 1.15.

Try It **Exploration A**

Throughout this chapter you will use the ε-δ definition of limit primarily to prove theorems about limits and to establish the existence or nonexistence of particular types of limits. For *finding* limits, you will learn techniques that are easier to use than the ε-δ definition of limit.

Section 1.3 Evaluating Limits Analytically

- **Evaluate a limit using properties of limits.**
- **Develop and use a strategy for finding limits.**
- **Evaluate a limit using dividing out and rationalizing techniques.**
- **Evaluate a limit using the Squeeze Theorem.**

Video

Properties of Limits

In Section 1.2, you learned that the limit of $f(x)$ as x approaches c does not depend on the value of f at $x = c$. It may happen, however, that the limit is precisely $f(c)$. In such cases, the limit can be evaluated by **direct substitution.** That is,

$$\lim_{x \to c} f(x) = f(c). \qquad \text{Substitute } c \text{ for } x.$$

Such *well-behaved* functions are **continuous at *c*.** You will examine this concept more closely in Section 1.4.

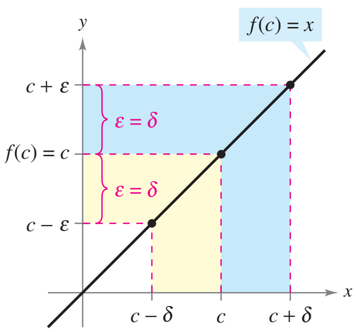


Figure 1.16

THEOREM 1.1 Some Basic Limits

Let b and c be real numbers and let n be a positive integer.

1. $\lim_{x \to c} b = b$ **2.** $\lim_{x \to c} x = c$ **3.** $\lim_{x \to c} x^n = c^n$

Proof To prove Property 2 of Theorem 1.1, you need to show that for each $\varepsilon > 0$ there exists a $\delta > 0$ such that $|x - c| < \varepsilon$ whenever $0 < |x - c| < \delta$. To do this, choose $\delta = \varepsilon$. The second inequality then implies the first, as shown in Figure 1.16. This completes the proof. (Proofs of the other properties of limits in this section are listed in Appendix A or are discussed in the exercises.)

NOTE When you encounter new notations or symbols in mathematics, be sure you know how the notations are read. For instance, the limit in Example 1(c) is read as "the limit of x^2 as x approaches 2 is 4."

EXAMPLE 1 Evaluating Basic Limits

a. $\lim_{x \to 2} 3 = 3$ **b.** $\lim_{x \to -4} x = -4$ **c.** $\lim_{x \to 2} x^2 = 2^2 = 4$

Try It Exploration A

The editable graph feature allows you to edit the graph of a function to visually evaluate the limit as x approaches c.

a. Editable Graph **b.** Editable Graph **c.** Editable Graph

THEOREM 1.2 Properties of Limits

Let b and c be real numbers, let n be a positive integer, and let f and g be functions with the following limits.

$$\lim_{x \to c} f(x) = L \quad \text{and} \quad \lim_{x \to c} g(x) = K$$

1. Scalar multiple: $\lim_{x \to c} [b f(x)] = bL$

2. Sum or difference: $\lim_{x \to c} [f(x) \pm g(x)] = L \pm K$

3. Product: $\lim_{x \to c} [f(x)g(x)] = LK$

4. Quotient: $\lim_{x \to c} \dfrac{f(x)}{g(x)} = \dfrac{L}{K}$, provided $K \neq 0$

5. Power: $\lim_{x \to c} [f(x)]^n = L^n$

EXAMPLE 2 The Limit of a Polynomial

$$\begin{aligned}\lim_{x\to 2}(4x^2+3) &= \lim_{x\to 2} 4x^2 + \lim_{x\to 2} 3 && \text{Property 2}\\ &= 4\left(\lim_{x\to 2} x^2\right) + \lim_{x\to 2} 3 && \text{Property 1}\\ &= 4(2^2)+3 && \text{Example 1}\\ &= 19 && \text{Simplify.}\end{aligned}$$

Try It

Exploration A

The editable graph feature allows you to edit the graph of a function to visually evaluate the limit as x approaches c.

Editable Graph

In Example 2, note that the limit (as $x \to 2$) of the *polynomial function* $p(x) = 4x^2 + 3$ is simply the value of p at $x = 2$.

$$\lim_{x\to 2} p(x) = p(2) = 4(2^2) + 3 = 19$$

This *direct substitution* property is valid for all polynomial and rational functions with nonzero denominators.

THEOREM 1.3 Limits of Polynomial and Rational Functions

If p is a polynomial function and c is a real number, then

$$\lim_{x\to c} p(x) = p(c).$$

If r is a rational function given by $r(x) = p(x)/q(x)$ and c is a real number such that $q(c) \neq 0$, then

$$\lim_{x\to c} r(x) = r(c) = \frac{p(c)}{q(c)}.$$

EXAMPLE 3 The Limit of a Rational Function

Find the limit: $\displaystyle\lim_{x\to 1} \frac{x^2+x+2}{x+1}$.

Solution Because the denominator is not 0 when $x = 1$, you can apply Theorem 1.3 to obtain

$$\lim_{x\to 1} \frac{x^2+x+2}{x+1} = \frac{1^2+1+2}{1+1} = \frac{4}{2} = 2.$$

Try It

Exploration A

The editable graph feature allows you to edit the graph of a function to visually evaluate the limit as x approaches c.

Editable Graph

Polynomial functions and rational functions are two of the three basic types of algebraic functions. The following theorem deals with the limit of the third type of algebraic function—one that involves a radical. See Appendix A for a proof of this theorem.

THEOREM 1.4 The Limit of a Function Involving a Radical

Let n be a positive integer. The following limit is valid for all c if n is odd, and is valid for $c > 0$ if n is even.

$$\lim_{x\to c} \sqrt[n]{x} = \sqrt[n]{c}$$

THE SQUARE ROOT SYMBOL

The first use of a symbol to denote the square root can be traced to the sixteenth century. Mathematicians first used the symbol $\surd$, which had only two strokes. This symbol was chosen because it resembled a lowercase r, to stand for the Latin word *radix*, meaning root.

Video

Video

The following theorem greatly expands your ability to evaluate limits because it shows how to analyze the limit of a composite function. See Appendix A for a proof of this theorem.

THEOREM 1.5 The Limit of a Composite Function

If f and g are functions such that $\lim_{x \to c} g(x) = L$ and $\lim_{x \to L} f(x) = f(L)$, then

$$\lim_{x \to c} f(g(x)) = f\left(\lim_{x \to c} g(x)\right) = f(L).$$

EXAMPLE 4 The Limit of a Composite Function

a. Because

$$\lim_{x \to 0} (x^2 + 4) = 0^2 + 4 = 4 \qquad \text{and} \qquad \lim_{x \to 4} \sqrt{x} = 2$$

it follows that

$$\lim_{x \to 0} \sqrt{x^2 + 4} = \sqrt{4} = 2.$$

b. Because

$$\lim_{x \to 3} (2x^2 - 10) = 2(3^2) - 10 = 8 \quad \text{and} \quad \lim_{x \to 8} \sqrt[3]{x} = 2$$

it follows that

$$\lim_{x \to 3} \sqrt[3]{2x^2 - 10} = \sqrt[3]{8} = 2.$$

Try It Exploration A Open Exploration

The editable graph feature allows you to edit the graph of a function to visually evaluate the limit as x approaches c.

a. Editable Graph **b.** Editable Graph

You have seen that the limits of many algebraic functions can be evaluated by direct substitution. The six basic trigonometric functions also exhibit this desirable quality, as shown in the next theorem (presented without proof).

THEOREM 1.6 Limits of Trigonometric Functions

Let c be a real number in the domain of the given trigonometric function.

1. $\lim_{x \to c} \sin x = \sin c$ **2.** $\lim_{x \to c} \cos x = \cos c$

3. $\lim_{x \to c} \tan x = \tan c$ **4.** $\lim_{x \to c} \cot x = \cot c$

5. $\lim_{x \to c} \sec x = \sec c$ **6.** $\lim_{x \to c} \csc x = \csc c$

EXAMPLE 5 Limits of Trigonometric Functions

a. $\lim_{x \to 0} \tan x = \tan(0) = 0$

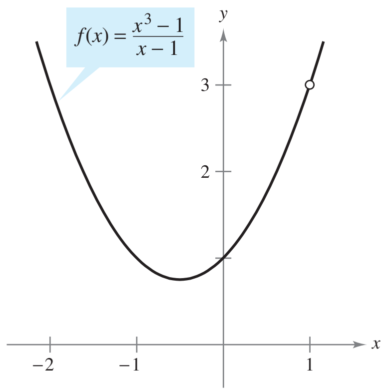
b. $\lim_{x \to \pi} (x \cos x) = \left(\lim_{x \to \pi} x\right)\left(\lim_{x \to \pi} \cos x\right) = \pi \cos(\pi) = -\pi$

c. $\lim_{x \to 0} \sin^2 x = \lim_{x \to 0} (\sin x)^2 = 0^2 = 0$

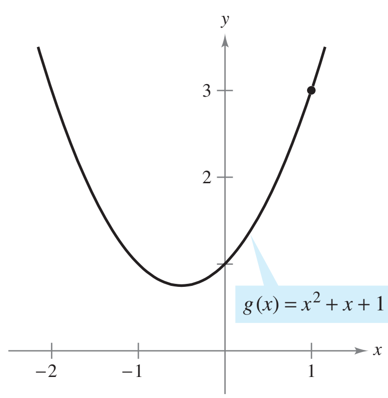
Try It Exploration A

A Strategy for Finding Limits

On the previous three pages, you studied several types of functions whose limits can be evaluated by direct substitution. This knowledge, together with the following theorem, can be used to develop a strategy for finding limits. A proof of this theorem is given in Appendix A.



Editable Graph



f and g agree at all but one point.

Editable Graph

Figure 1.17

THEOREM 1.7 Functions That Agree at All But One Point

Let c be a real number and let $f(x) = g(x)$ for all $x \neq c$ in an open interval containing c. If the limit of $g(x)$ as x approaches c exists, then the limit of $f(x)$ also exists and

$$\lim_{x \to c} f(x) = \lim_{x \to c} g(x).$$

EXAMPLE 6 Finding the Limit of a Function

Find the limit: $\lim_{x \to 1} \frac{x^3 - 1}{x - 1}$.

Solution Let $f(x) = (x^3 - 1)/(x - 1)$. By factoring and dividing out like factors, you can rewrite f as

$$f(x) = \frac{\cancel{(x - 1)}(x^2 + x + 1)}{\cancel{(x - 1)}} = x^2 + x + 1 = g(x), \qquad x \neq 1.$$

So, for all x-values other than $x = 1$, the functions f and g agree, as shown in Figure 1.17. Because $\lim_{x \to 1} g(x)$ exists, you can apply Theorem 1.7 to conclude that f and g have the same limit at $x = 1$.

$$\begin{aligned}
\lim_{x \to 1} \frac{x^3 - 1}{x - 1} &= \lim_{x \to 1} \frac{(x - 1)(x^2 + x + 1)}{x - 1} && \text{Factor.} \\
&= \lim_{x \to 1} \frac{\cancel{(x - 1)}(x^2 + x + 1)}{\cancel{x - 1}} && \text{Divide out like factors.} \\
&= \lim_{x \to 1} (x^2 + x + 1) && \text{Apply Theorem 1.7.} \\
&= 1^2 + 1 + 1 && \text{Use direct substitution.} \\
&= 3 && \text{Simplify.}
\end{aligned}$$

Try It | Exploration A | Exploration B | Exploration C | Exploration D

STUDY TIP When applying this strategy for finding a limit, remember that some functions do not have a limit (as x approaches c). For instance, the following limit does not exist.

$$\lim_{x \to 1} \frac{x^3 + 1}{x - 1}$$

A Strategy for Finding Limits

1. Learn to recognize which limits can be evaluated by direct substitution. (These limits are listed in Theorems 1.1 through 1.6.)
2. If the limit of $f(x)$ as x approaches c *cannot* be evaluated by direct substitution, try to find a function g that agrees with f for all x other than $x = c$. [Choose g such that the limit of $g(x)$ *can* be evaluated by direct substitution.]
3. Apply Theorem 1.7 to conclude *analytically* that

$$\lim_{x \to c} f(x) = \lim_{x \to c} g(x) = g(c).$$

4. Use a *graph* or *table* to reinforce your conclusion.

Dividing Out and Rationalizing Techniques

Two techniques for finding limits analytically are shown in Examples 7 and 8. The first technique involves dividing out common factors, and the second technique involves rationalizing the numerator of a fractional expression.

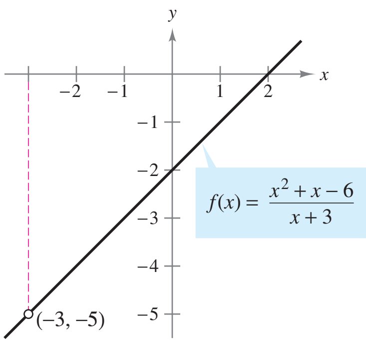
EXAMPLE 7 Dividing Out Technique

Find the limit: $\lim\limits_{x \to -3} \dfrac{x^2 + x - 6}{x + 3}$.

Solution Although you are taking the limit of a rational function, you *cannot* apply Theorem 1.3 because the limit of the denominator is 0.

$$\lim_{x \to -3} \frac{x^2 + x - 6}{x + 3} \quad \begin{cases} \lim\limits_{x \to -3} (x^2 + x - 6) = 0 \\ \lim\limits_{x \to -3} (x + 3) = 0 \end{cases}$$

Direct substitution fails.



f is undefined when $x = -3$.
Figure 1.18

Editable Graph

Because the limit of the numerator is also 0, the numerator and denominator have a *common factor* of $(x + 3)$. So, for all $x \neq -3$, you can divide out this factor to obtain

$$f(x) = \frac{x^2 + x - 6}{x + 3} = \frac{\cancel{(x + 3)}(x - 2)}{\cancel{x + 3}} = x - 2 = g(x), \qquad x \neq -3.$$

Using Theorem 1.7, it follows that

$$\lim_{x \to -3} \frac{x^2 + x - 6}{x + 3} = \lim_{x \to -3} (x - 2) \qquad \text{Apply Theorem 1.7.}$$

$$= -5. \qquad \text{Use direct substitution.}$$

This result is shown graphically in Figure 1.18. Note that the graph of the function f coincides with the graph of the function $g(x) = x - 2$, except that the graph of f has a gap at the point $(-3, -5)$.

Try It | Exploration A | Open Exploration

NOTE In the solution of Example 7, be sure you see the usefulness of the Factor Theorem of Algebra. This theorem states that if c is a zero of a polynomial function, $(x - c)$ is a factor of the polynomial. So, if you apply direct substitution to a rational function and obtain

$$r(c) = \frac{p(c)}{q(c)} = \frac{0}{0}$$

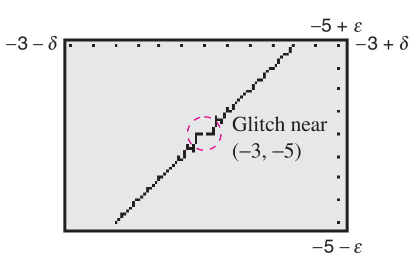
you can conclude that $(x - c)$ must be a common factor to both $p(x)$ and $q(x)$.

In Example 7, direct substitution produced the meaningless fractional form 0/0. An expression such as 0/0 is called an **indeterminate form** because you cannot (from the form alone) determine the limit. When you try to evaluate a limit and encounter this form, remember that you must rewrite the fraction so that the new denominator does not have 0 as its limit. One way to do this is to *divide out like factors*, as shown in Example 7. A second way is to *rationalize the numerator*, as shown in Example 8.

TECHNOLOGY PITFALL Because the graphs of

$$f(x) = \frac{x^2 + x - 6}{x + 3} \qquad \text{and} \qquad g(x) = x - 2$$

differ only at the point $(-3, -5)$, a standard graphing utility setting may not distinguish clearly between these graphs. However, because of the pixel configuration and rounding error of a graphing utility, it may be possible to find screen settings that distinguish between the graphs. Specifically, by repeatedly zooming in near the point $(-3, -5)$ on the graph of f, your graphing utility may show glitches or irregularities that do not exist on the actual graph. (See Figure 1.19.) By changing the screen settings on your graphing utility you may obtain the correct graph of f.



Incorrect graph of f
Figure 1.19

EXAMPLE 8 Rationalizing Technique

Find the limit: $\lim\limits_{x\to 0} \dfrac{\sqrt{x+1}-1}{x}$.

Solution By direct substitution, you obtain the indeterminate form $0/0$.

$$\lim_{x\to 0} \frac{\sqrt{x+1}-1}{x} \quad \begin{cases} \lim\limits_{x\to 0} \left(\sqrt{x+1}-1\right) = 0 \\ \lim\limits_{x\to 0} x = 0 \end{cases}$$

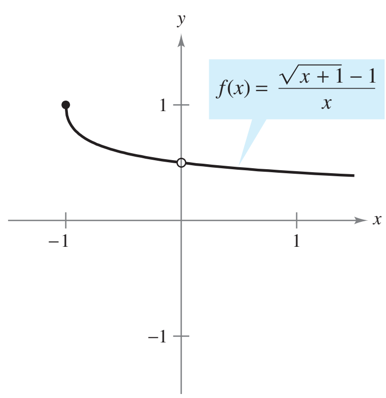
Direct substitution fails.

In this case, you can rewrite the fraction by rationalizing the numerator.

$$\begin{aligned}
\frac{\sqrt{x+1}-1}{x} &= \left(\frac{\sqrt{x+1}-1}{x}\right)\left(\frac{\sqrt{x+1}+1}{\sqrt{x+1}+1}\right) \\
&= \frac{(x+1)-1}{x\left(\sqrt{x+1}+1\right)} \\
&= \frac{\cancel{x}}{\cancel{x}\left(\sqrt{x+1}+1\right)} \\
&= \frac{1}{\sqrt{x+1}+1}, \quad x \neq 0
\end{aligned}$$

Now, using Theorem 1.7, you can evaluate the limit as shown.

$$\begin{aligned}
\lim_{x\to 0} \frac{\sqrt{x+1}-1}{x} &= \lim_{x\to 0} \frac{1}{\sqrt{x+1}+1} \\
&= \frac{1}{1+1} \\
&= \frac{1}{2}
\end{aligned}$$



The limit of $f(x)$ as x approaches 0 is $\frac{1}{2}$.
Figure 1.20

Editable Graph

A table or a graph can reinforce your conclusion that the limit is $\frac{1}{2}$. (See Figure 1.20.)

x approaches 0 from the left. → ← x approaches 0 from the right.

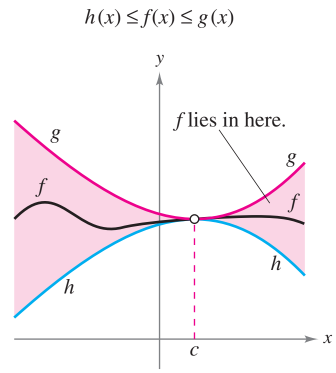
x	-0.25	-0.1	-0.01	-0.001	0	0.001	0.01	0.1	0.25
$f(x)$	0.5359	0.5132	0.5013	0.5001	?	0.4999	0.4988	0.4881	0.4721

$f(x)$ approaches 0.5. → ← $f(x)$ approaches 0.5.

Try It | Exploration A | Exploration B | Exploration C

NOTE The rationalizing technique for evaluating limits is based on multiplication by a convenient form of 1. In Example 8, the convenient form is

$$1 = \frac{\sqrt{x+1}+1}{\sqrt{x+1}+1}.$$



The Squeeze Theorem
Figure 1.21

The Squeeze Theorem

The next theorem concerns the limit of a function that is squeezed between two other functions, each of which has the same limit at a given x-value, as shown in Figure 1.21. (The proof of this theorem is given in Appendix A.)

THEOREM 1.8 The Squeeze Theorem

If $h(x) \le f(x) \le g(x)$ for all x in an open interval containing c, except possibly at c itself, and if

$$\lim_{x\to c} h(x) = L = \lim_{x\to c} g(x)$$

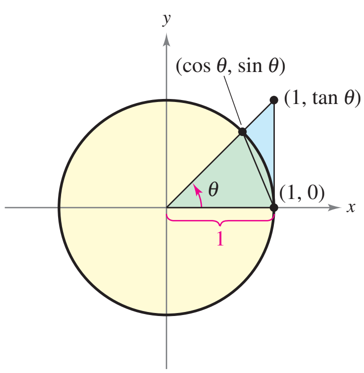
then $\lim_{x\to c} f(x)$ exists and is equal to L.

Video

You can see the usefulness of the Squeeze Theorem in the proof of Theorem 1.9.

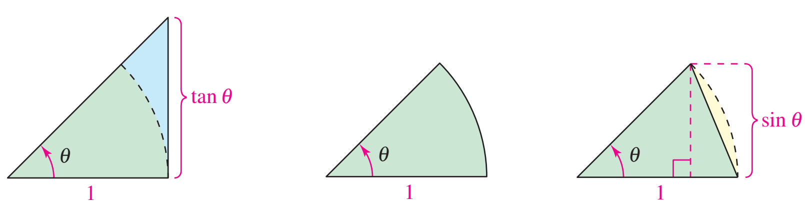
THEOREM 1.9 Two Special Trigonometric Limits

1. $\displaystyle \lim_{x\to 0} \frac{\sin x}{x} = 1$ **2.** $\displaystyle \lim_{x\to 0} \frac{1 - \cos x}{x} = 0$



A circular sector is used to prove Theorem 1.9.
Figure 1.22

Proof To avoid the confusion of two different uses of x, the proof is presented using the variable θ, where θ is an acute positive angle *measured in radians*. Figure 1.22 shows a circular sector that is squeezed between two triangles.



Area of triangle	$\ge$	Area of sector	$\ge$	Area of triangle
$\dfrac{\tan\theta}{2}$	$\ge$	$\dfrac{\theta}{2}$	$\ge$	$\dfrac{\sin\theta}{2}$

Multiplying each expression by $2/\sin\theta$ produces

$$\frac{1}{\cos\theta} \ge \frac{\theta}{\sin\theta} \ge 1$$

and taking reciprocals and reversing the inequalities yields

$$\cos\theta \le \frac{\sin\theta}{\theta} \le 1.$$

Because $\cos\theta = \cos(-\theta)$ and $(\sin\theta)/\theta = [\sin(-\theta)]/(-\theta)$, you can conclude that this inequality is valid for *all* nonzero θ in the open interval $(-\pi/2, \pi/2)$. Finally, because $\lim_{\theta\to 0} \cos\theta = 1$ and $\lim_{\theta\to 0} 1 = 1$, you can apply the Squeeze Theorem to conclude that $\lim_{\theta\to 0} (\sin\theta)/\theta = 1$. The proof of the second limit is left as an exercise (see Exercise 120).

FOR FURTHER INFORMATION For more information on the function $f(x) = (\sin x)/x$, see the article "The Function $(\sin x)/x$" by William B. Gearhart and Harris S. Shultz in *The College Mathematics Journal*.

MathArticle

EXAMPLE 9 A Limit Involving a Trigonometric Function

Find the limit: $\lim_{x \to 0} \frac{\tan x}{x}$.

Solution Direct substitution yields the indeterminate form 0/0. To solve this problem, you can write $\tan x$ as $(\sin x)/(\cos x)$ and obtain

$$\lim_{x \to 0} \frac{\tan x}{x} = \lim_{x \to 0} \left(\frac{\sin x}{x}\right)\left(\frac{1}{\cos x}\right).$$

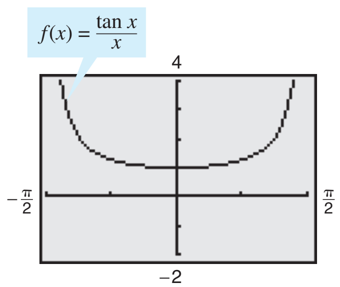
Now, because

$$\lim_{x \to 0} \frac{\sin x}{x} = 1 \qquad \text{and} \qquad \lim_{x \to 0} \frac{1}{\cos x} = 1$$

you can obtain

$$\begin{aligned} \lim_{x \to 0} \frac{\tan x}{x} &= \left(\lim_{x \to 0} \frac{\sin x}{x}\right)\left(\lim_{x \to 0} \frac{1}{\cos x}\right) \\ &= (1)(1) \\ &= 1. \end{aligned}$$

(See Figure 1.23.)



The limit of $f(x)$ as x approaches 0 is 1.
Figure 1.23

Editable Graph

Try It

Exploration A

EXAMPLE 10 A Limit Involving a Trigonometric Function

Find the limit: $\lim_{x \to 0} \frac{\sin 4x}{x}$.

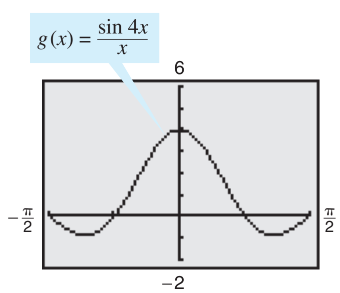
Solution Direct substitution yields the indeterminate form 0/0. To solve this problem, you can rewrite the limit as

$$\lim_{x \to 0} \frac{\sin 4x}{x} = 4\left(\lim_{x \to 0} \frac{\sin 4x}{4x}\right). \qquad \text{Multiply and divide by 4.}$$

Now, by letting $y = 4x$ and observing that $x \to 0$ if and only if $y \to 0$, you can write

$$\begin{aligned} \lim_{x \to 0} \frac{\sin 4x}{x} &= 4\left(\lim_{x \to 0} \frac{\sin 4x}{4x}\right) \\ &= 4\left(\lim_{y \to 0} \frac{\sin y}{y}\right) \\ &= 4(1) \\ &= 4. \end{aligned}$$

(See Figure 1.24.)



The limit of $g(x)$ as x approaches 0 is 4.
Figure 1.24

Editable Graph

Try It

Exploration A

TECHNOLOGY Use a graphing utility to confirm the limits in the examples and exercise set. For instance, Figures 1.23 and 1.24 show the graphs of

$$f(x) = \frac{\tan x}{x} \qquad \text{and} \qquad g(x) = \frac{\sin 4x}{x}.$$

Note that the first graph appears to contain the point (0, 1) and the second graph appears to contain the point (0, 4), which lends support to the conclusions obtained in Examples 9 and 10.

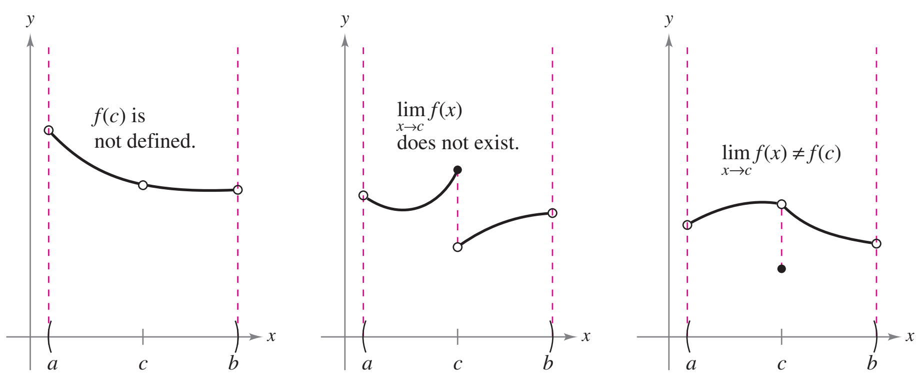
Section 1.4 Continuity and One-Sided Limits

- **Determine continuity at a point and continuity on an open interval.**
- **Determine one-sided limits and continuity on a closed interval.**
- **Use properties of continuity.**
- **Understand and use the Intermediate Value Theorem.**

Continuity at a Point and on an Open Interval

In mathematics, the term *continuous* has much the same meaning as it has in everyday usage. Informally, to say that a function f is continuous at $x = c$ means that there is no interruption in the graph of f at c. That is, its graph is unbroken at c and there are no holes, jumps, or gaps. Figure 1.25 identifies three values of x at which the graph of f is *not* continuous. At all other points in the interval (a, b), the graph of f is uninterrupted and **continuous.**

Animation



Three conditions exist for which the graph of f is not continuous at $x = c$.
Figure 1.25

EXPLORATION

Informally, you might say that a function is *continuous* on an open interval if its graph can be drawn with a pencil without lifting the pencil from the paper. Use a graphing utility to graph each function on the given interval. From the graphs, which functions would you say are continuous on the interval? Do you think you can trust the results you obtained graphically? Explain your reasoning.

Function	*Interval*
a. $y = x^2 + 1$	$(-3, 3)$
b. $y = \dfrac{1}{x-2}$	$(-3, 3)$
c. $y = \dfrac{\sin x}{x}$	$(-\pi, \pi)$
d. $y = \dfrac{x^2 - 4}{x + 2}$	$(-3, 3)$
e. $y = \begin{cases} 2x - 4, & x \le 0 \\ x + 1, & x > 0 \end{cases}$	$(-3, 3)$

In Figure 1.25, it appears that continuity at $x = c$ can be destroyed by any one of the following conditions.

1. The function is not defined at $x = c$.
2. The limit of $f(x)$ does not exist at $x = c$.
3. The limit of $f(x)$ exists at $x = c$, but it is not equal to $f(c)$.

If *none* of the above three conditions is true, the function f is called **continuous at c,** as indicated in the following important definition.

Definition of Continuity

Continuity at a Point: A function f is **continuous at c** if the following three conditions are met.

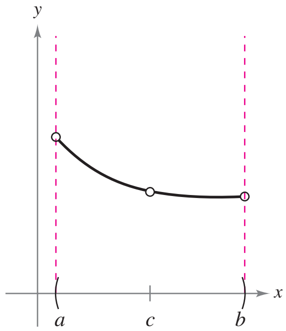
1. $f(c)$ is defined.
2. $\lim_{x\to c} f(x)$ exists.
3. $\lim_{x\to c} f(x) = f(c)$.

Continuity on an Open Interval: A function is **continuous on an open interval (a, b)** if it is continuous at each point in the interval. A function that is continuous on the entire real line $(-\infty, \infty)$ is **everywhere continuous.**

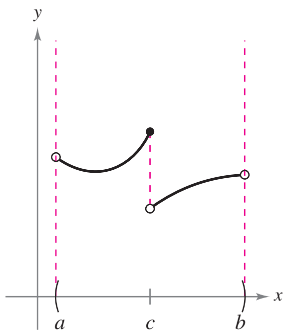
Video

FOR FURTHER INFORMATION For more information on the concept of continuity, see the article "Leibniz and the Spell of the Continuous" by Hardy Grant in *The College Mathematics Journal*.

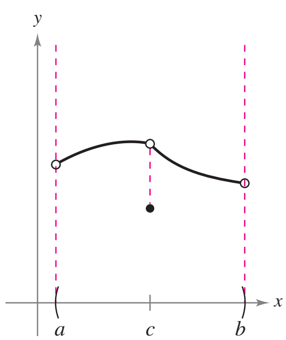
MathArticle



(a) Removable discontinuity



(b) Nonremovable discontinuity



(c) Removable discontinuity

Figure 1.26

Consider an open interval I that contains a real number c. If a function f is defined on I (except possibly at c), and f is not continuous at c, then f is said to have a **discontinuity** at c. Discontinuities fall into two categories: **removable** and **nonremovable.** A discontinuity at c is called removable if f can be made continuous by appropriately defining (or redefining) $f(c)$. For instance, the functions shown in Figure 1.26(a) and (c) have removable discontinuities at c and the function shown in Figure 1.26(b) has a nonremovable discontinuity at c.

EXAMPLE 1 Continuity of a Function

Discuss the continuity of each function.

a. $f(x) = \dfrac{1}{x}$ **b.** $g(x) = \dfrac{x^2 - 1}{x - 1}$ **c.** $h(x) = \begin{cases} x + 1, & x \le 0 \\ x^2 + 1, & x > 0 \end{cases}$ **d.** $y = \sin x$

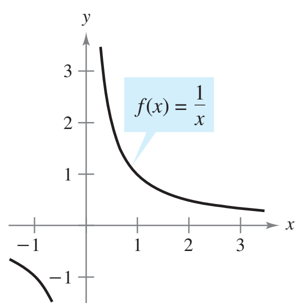
Solution

a. The domain of f is all nonzero real numbers. From Theorem 1.3, you can conclude that f is continuous at every x-value in its domain. At $x = 0$, f has a nonremovable discontinuity, as shown in Figure 1.27(a). In other words, there is no way to define $f(0)$ so as to make the function continuous at $x = 0$.

b. The domain of g is all real numbers except $x = 1$. From Theorem 1.3, you can conclude that g is continuous at every x-value in its domain. At $x = 1$, the function has a removable discontinuity, as shown in Figure 1.27(b). If $g(1)$ is defined as 2, the "newly defined" function is continuous for all real numbers.

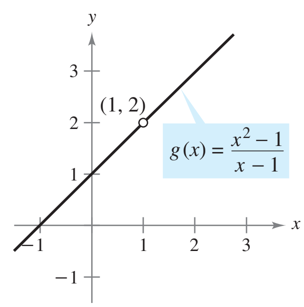
c. The domain of h is all real numbers. The function h is continuous on $(-\infty, 0)$ and $(0, \infty)$, and, because $\lim_{x \to 0} h(x) = 1$, h is continuous on the entire real line, as shown in Figure 1.27(c).

d. The domain of y is all real numbers. From Theorem 1.6, you can conclude that the function is continuous on its entire domain, $(-\infty, \infty)$, as shown in Figure 1.27(d).



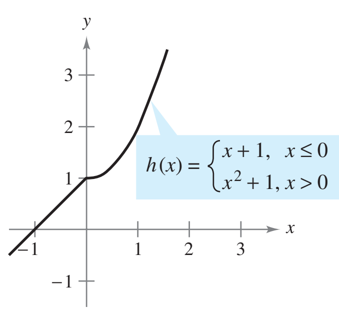
(a) Nonremovable discontinuity at $x = 0$

Editable Graph



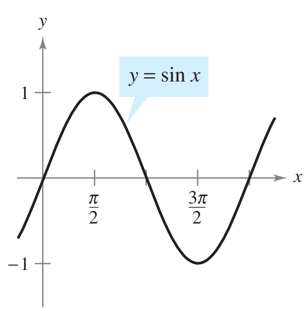
(b) Removable discontinuity at $x = 1$

Editable Graph



(c) Continuous on entire real line

Editable Graph



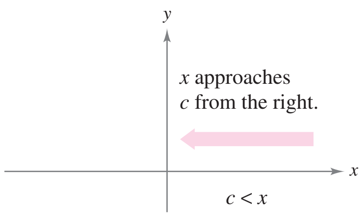
(d) Continuous on entire real line

Editable Graph

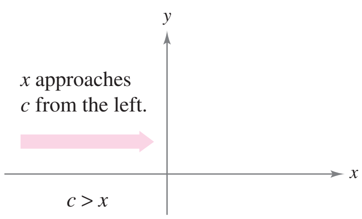
Figure 1.27

STUDY TIP Some people may refer to the function in Example 1(a) as "discontinuous." We have found that this terminology can be confusing. Rather than saying the function is discontinuous, we prefer to say that it has a discontinuity at $x = 0$.

Try It Exploration A Exploration B Exploration C



(a) Limit from right



(b) Limit from left
Figure 1.28

One-Sided Limits and Continuity on a Closed Interval

To understand continuity on a closed interval, you first need to look at a different type of limit called a **one-sided limit.** For example, the **limit from the right** means that x approaches c from values greater than c [see Figure 1.28(a)]. This limit is denoted as

$$\lim_{x \to c^+} f(x) = L. \qquad \text{Limit from the right}$$

Similarly, the **limit from the left** means that x approaches c from values less than c [see Figure 1.28(b)]. This limit is denoted as

$$\lim_{x \to c^-} f(x) = L. \qquad \text{Limit from the left}$$

One-sided limits are useful in taking limits of functions involving radicals. For instance, if n is an even integer,

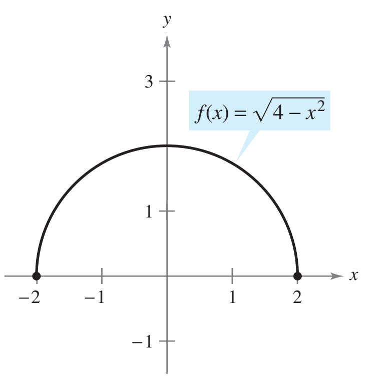
$$\lim_{x \to 0^+} \sqrt[n]{x} = 0.$$

EXAMPLE 2 A One-Sided Limit

Find the limit of $f(x) = \sqrt{4 - x^2}$ as x approaches -2 from the right.

Solution As shown in Figure 1.29, the limit as x approaches -2 from the right is

$$\lim_{x \to -2^+} \sqrt{4 - x^2} = 0.$$



The limit of $f(x)$ as x approaches -2 from the right is 0.
Figure 1.29

Editable Graph

Try It

Exploration A

One-sided limits can be used to investigate the behavior of **step functions.** One common type of step function is the **greatest integer function** $[\![x]\!]$, defined by

$$[\![x]\!] = \text{greatest integer } n \text{ such that } n \le x. \qquad \text{Greatest integer function}$$

For instance, $[\![2.5]\!] = 2$ and $[\![-2.5]\!] = -3$.

EXAMPLE 3 The Greatest Integer Function

Find the limit of the greatest integer function $f(x) = [\![x]\!]$ as x approaches 0 from the left and from the right.

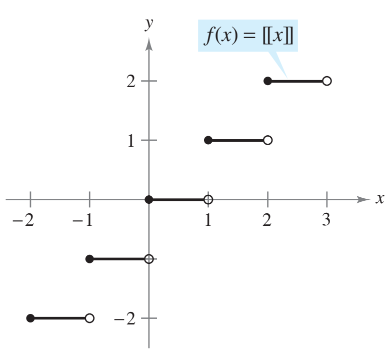
Solution As shown in Figure 1.30, the limit as x approaches 0 *from the left* is given by

$$\lim_{x \to 0^-} [\![x]\!] = -1$$

and the limit as x approaches 0 *from the right* is given by

$$\lim_{x \to 0^+} [\![x]\!] = 0.$$

The greatest integer function has a discontinuity at zero because the left and right limits at zero are different. By similar reasoning, you can see that the greatest integer function has a discontinuity at any integer n.



Greatest integer function
Figure 1.30

Editable Graph

Try It

Exploration A

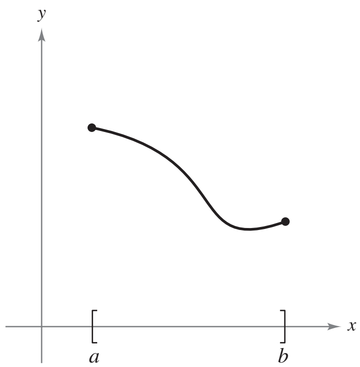
Exploration B

When the limit from the left is not equal to the limit from the right, the (two-sided) limit *does not exist*. The next theorem makes this more explicit. The proof of this theorem follows directly from the definition of a one-sided limit.

THEOREM 1.10 The Existence of a Limit

Let f be a function and let c and L be real numbers. The limit of $f(x)$ as x approaches c is L if and only if

$$\lim_{x\to c^-} f(x) = L \quad \text{and} \quad \lim_{x\to c^+} f(x) = L.$$



Continuous function on a closed interval
Figure 1.31

The concept of a one-sided limit allows you to extend the definition of continuity to closed intervals. Basically, a function is continuous on a closed interval if it is continuous in the interior of the interval and exhibits one-sided continuity at the endpoints. This is stated formally as follows.

Definition of Continuity on a Closed Interval

A function f is **continuous on the closed interval $[a, b]$** if it is continuous on the open interval (a, b) and

$$\lim_{x\to a^+} f(x) = f(a) \quad \text{and} \quad \lim_{x\to b^-} f(x) = f(b).$$

The function f is **continuous from the right** at a and **continuous from the left** at b (see Figure 1.31).

Similar definitions can be made to cover continuity on intervals of the form $(a, b]$ and $[a, b)$ that are neither open nor closed, or on infinite intervals. For example, the function

$$f(x) = \sqrt{x}$$

is continuous on the infinite interval $[0, \infty)$, and the function

$$g(x) = \sqrt{2 - x}$$

is continuous on the infinite interval $(-\infty, 2]$.

EXAMPLE 4 Continuity on a Closed Interval

Discuss the continuity of $f(x) = \sqrt{1 - x^2}$.

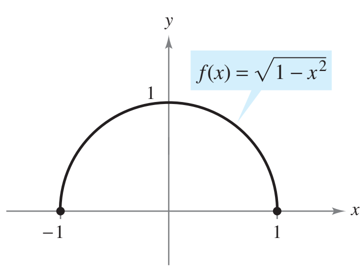
Solution The domain of f is the closed interval $[-1, 1]$. At all points in the open interval $(-1, 1)$, the continuity of f follows from Theorems 1.4 and 1.5. Moreover, because

$$\lim_{x\to -1^+} \sqrt{1 - x^2} = 0 = f(-1) \qquad \text{Continuous from the right}$$

and

$$\lim_{x\to 1^-} \sqrt{1 - x^2} = 0 = f(1) \qquad \text{Continuous from the left}$$

you can conclude that f is continuous on the closed interval $[-1, 1]$, as shown in Figure 1.32.



f is continuous on $[-1, 1]$.
Figure 1.32

Editable Graph

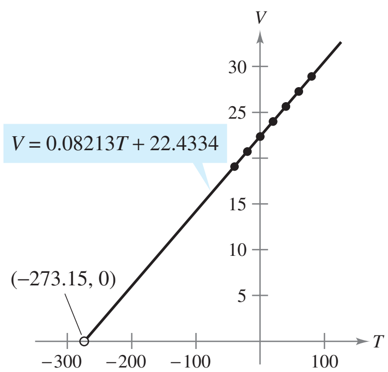
Try It

Exploration A

The next example shows how a one-sided limit can be used to determine the value of absolute zero on the Kelvin scale.

EXAMPLE 5 Charles's Law and Absolute Zero

On the Kelvin scale, *absolute zero* is the temperature 0 K. Although temperatures of approximately 0.0001 K have been produced in laboratories, absolute zero has never been attained. In fact, evidence suggests that absolute zero *cannot* be attained. How did scientists determine that 0 K is the "lower limit" of the temperature of matter? What is absolute zero on the Celsius scale?



The volume of hydrogen gas depends on its temperature.
Figure 1.33

Editable Graph

Solution The determination of absolute zero stems from the work of the French physicist Jacques Charles (1746–1823). Charles discovered that the volume of gas at a constant pressure increases linearly with the temperature of the gas. The table illustrates this relationship between volume and temperature. In the table, one mole of hydrogen is held at a constant pressure of one atmosphere. The volume V is measured in liters and the temperature T is measured in degrees Celsius.

T	-40	-20	0	20	40	60	80
V	19.1482	20.7908	22.4334	24.0760	25.7186	27.3612	29.0038

The points represented by the table are shown in Figure 1.33. Moreover, by using the points in the table, you can determine that T and V are related by the linear equation

$$V = 0.08213T + 22.4334 \qquad \text{or} \qquad T = \frac{V - 22.4334}{0.08213}.$$

By reasoning that the volume of the gas can approach 0 (but never equal or go below 0) you can determine that the "least possible temperature" is given by

$$\begin{aligned} \lim_{V\to 0^+} T &= \lim_{V\to 0^+} \frac{V - 22.4334}{0.08213} \\ &= \frac{0 - 22.4334}{0.08213} && \text{Use direct substitution.} \\ &\approx -273.15. \end{aligned}$$

So, absolute zero on the Kelvin scale (0 K) is approximately $-273.15°$ on the Celsius scale.

Try It   Exploration A

The following table shows the temperatures in Example 5, converted to the Fahrenheit scale. Try repeating the solution shown in Example 5 using these temperatures and volumes. Use the result to find the value of absolute zero on the Fahrenheit scale.

T	-40	-4	32	68	104	140	176
V	19.1482	20.7908	22.4334	24.0760	25.7186	27.3612	29.0038

In 1995, physicists Carl Wieman and Eric Cornell of the University of Colorado at Boulder used lasers and evaporation to produce a supercold gas in which atoms overlap. This gas is called a Bose-Einstein condensate. "We get to within a billionth of a degree of absolute zero," reported Wieman. (*Source:* Time *magazine, April 10, 2000*)

NOTE Charles's Law for gases (assuming constant pressure) can be stated as

$$V = RT \qquad \text{Charles's Law}$$

where V is volume, R is constant, and T is temperature. In the statement of this law, what property must the temperature scale have?

AUGUSTIN-LOUIS CAUCHY (1789–1857)

The concept of a continuous function was first introduced by Augustin-Louis Cauchy in 1821. The definition given in his text *Cours d'Analyse* stated that indefinite small changes in y were the result of indefinite small changes in x. "… $f(x)$ will be called a *continuous* function if … the numerical values of the difference $f(x + \alpha) - f(x)$ decrease indefinitely with those of α …."

MathBio

Properties of Continuity

In Section 1.3, you studied several properties of limits. Each of those properties yields a corresponding property pertaining to the continuity of a function. For instance, Theorem 1.11 follows directly from Theorem 1.2.

THEOREM 1.11 Properties of Continuity

If b is a real number and f and g are continuous at $x = c$, then the following functions are also continuous at c.

1. Scalar multiple: bf
2. Sum and difference: $f \pm g$
3. Product: fg
4. Quotient: $\dfrac{f}{g}$, if $g(c) \neq 0$

The following types of functions are continuous at every point in their domains.

1. Polynomial functions: $p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdot \cdot \cdot + a_1 x + a_0$
2. Rational functions: $r(x) = \dfrac{p(x)}{q(x)}, \quad q(x) \neq 0$
3. Radical functions: $f(x) = \sqrt[n]{x}$
4. Trigonometric functions: $\sin x$, $\cos x$, $\tan x$, $\cot x$, $\sec x$, $\csc x$

By combining Theorem 1.11 with this summary, you can conclude that a wide variety of elementary functions are continuous at every point in their domains.

EXAMPLE 6 Applying Properties of Continuity

By Theorem 1.11, it follows that each of the following functions is continuous at every point in its domain.

$$f(x) = x + \sin x, \qquad f(x) = 3 \tan x, \qquad f(x) = \frac{x^2 + 1}{\cos x}$$

Try It | Exploration A | Open Exploration

The next theorem, which is a consequence of Theorem 1.5, allows you to determine the continuity of *composite* functions such as

$$f(x) = \sin 3x, \qquad f(x) = \sqrt{x^2 + 1}, \qquad f(x) = \tan \frac{1}{x}.$$

THEOREM 1.12 Continuity of a Composite Function

If g is continuous at c and f is continuous at $g(c)$, then the composite function given by $(f \circ g)(x) = f(g(x))$ is continuous at c.

One consequence of Theorem 1.12 is that if f and g satisfy the given conditions, you can determine the limit of $f(g(x))$ as x approaches c to be

$$\lim_{x \to c} f(g(x)) = f(g(c)).$$

Technology

EXAMPLE 7 Testing for Continuity

Describe the interval(s) on which each function is continuous.

a. $f(x) = \tan x$ **b.** $g(x) = \begin{cases} \sin \dfrac{1}{x}, & x \neq 0 \\ 0, & x = 0 \end{cases}$ **c.** $h(x) = \begin{cases} x \sin \dfrac{1}{x}, & x \neq 0 \\ 0, & x = 0 \end{cases}$

Solution

a. The tangent function $f(x) = \tan x$ is undefined at

$$x = \frac{\pi}{2} + n\pi, \qquad n \text{ is an integer.}$$

At all other points it is continuous. So, $f(x) = \tan x$ is continuous on the open intervals

$$\ldots, \left(-\frac{3\pi}{2}, -\frac{\pi}{2}\right), \left(-\frac{\pi}{2}, \frac{\pi}{2}\right), \left(\frac{\pi}{2}, \frac{3\pi}{2}\right), \ldots$$

as shown in Figure 1.34(a).

b. Because $y = 1/x$ is continuous except at $x = 0$ and the sine function is continuous for all real values of x, it follows that $y = \sin(1/x)$ is continuous at all real values except $x = 0$. At $x = 0$, the limit of $g(x)$ does not exist (see Example 5, Section 1.2). So, g is continuous on the intervals $(-\infty, 0)$ and $(0, \infty)$, as shown in Figure 1.34(b).

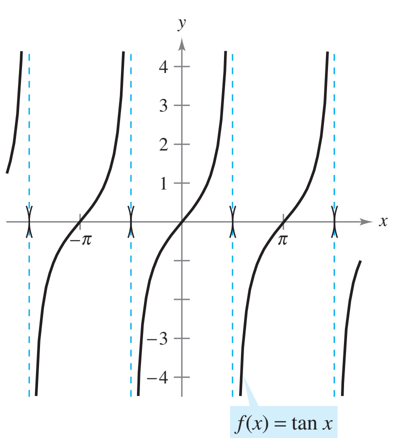
c. This function is similar to that in part (b) except that the oscillations are damped by the factor x. Using the Squeeze Theorem, you obtain

$$-|x| \leq x \sin \frac{1}{x} \leq |x|, \qquad x \neq 0$$

and you can conclude that

$$\lim_{x \to 0} h(x) = 0.$$

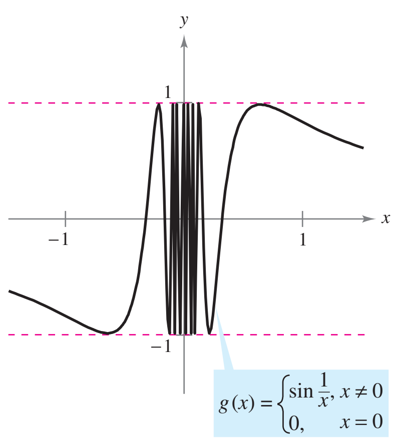
So, h is continuous on the entire real line, as shown in Figure 1.34(c).



(a) f is continuous on each open interval in its domain.

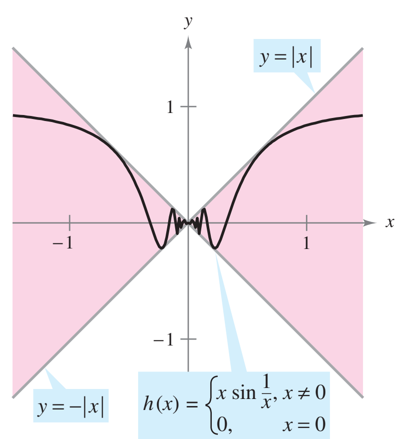
Editable Graph

Figure 1.34



(b) g is continuous on $(-\infty, 0)$ and $(0, \infty)$.

Editable Graph



(c) h is continuous on the entire real line.

Editable Graph

Try It

Exploration A

The Intermediate Value Theorem

Theorem 1.13 is an important theorem concerning the behavior of functions that are continuous on a closed interval.

THEOREM 1.13 Intermediate Value Theorem

If f is continuous on the closed interval $[a, b]$ and k is any number between $f(a)$ and $f(b)$, then there is at least one number c in $[a, b]$ such that

$$f(c) = k.$$

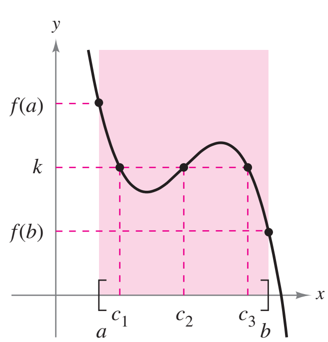
Video

NOTE The Intermediate Value Theorem tells you that at least one c exists, but it does not give a method for finding c. Such theorems are called **existence theorems.**

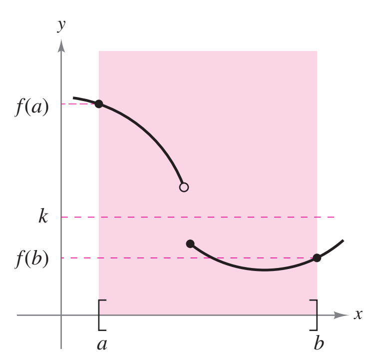
By referring to a text on advanced calculus, you will find that a proof of this theorem is based on a property of real numbers called *completeness*. The Intermediate Value Theorem states that for a continuous function f, if x takes on all values between a and b, $f(x)$ must take on all values between $f(a)$ and $f(b)$.

As a simple example of this theorem, consider a person's height. Suppose that a girl is 5 feet tall on her thirteenth birthday and 5 feet 7 inches tall on her fourteenth birthday. Then, for any height h between 5 feet and 5 feet 7 inches, there must have been a time t when her height was exactly h. This seems reasonable because human growth is continuous and a person's height does not abruptly change from one value to another.

The Intermediate Value Theorem guarantees the existence of *at least one* number c in the closed interval $[a, b]$. There may, of course, be more than one number c such that $f(c) = k$, as shown in Figure 1.35. A function that is not continuous does not necessarily exhibit the intermediate value property. For example, the graph of the function shown in Figure 1.36 jumps over the horizontal line given by $y = k$, and for this function there is no value of c in $[a, b]$ such that $f(c) = k$.

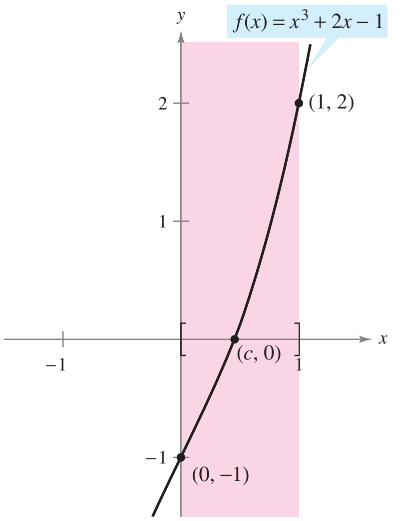


f is continuous on $[a, b]$.
[There exist three c's such that $f(c) = k$.]
Figure 1.35



f is not continuous on $[a, b]$.
[There are no c's such that $f(c) = k$.]
Figure 1.36

The Intermediate Value Theorem often can be used to locate the zeros of a function that is continuous on a closed interval. Specifically, if f is continuous on $[a, b]$ and $f(a)$ and $f(b)$ differ in sign, the Intermediate Value Theorem guarantees the existence of at least one zero of f in the closed interval $[a, b]$.



f is continuous on $[0, 1]$ with $f(0) < 0$ and $f(1) > 0$.
Figure 1.37

Editable Graph

EXAMPLE 8 An Application of the Intermediate Value Theorem

Use the Intermediate Value Theorem to show that the polynomial function $f(x) = x^3 + 2x - 1$ has a zero in the interval $[0, 1]$.

Solution Note that f is continuous on the closed interval $[0, 1]$. Because

$$f(0) = 0^3 + 2(0) - 1 = -1 \quad \text{and} \quad f(1) = 1^3 + 2(1) - 1 = 2$$

it follows that $f(0) < 0$ and $f(1) > 0$. You can therefore apply the Intermediate Value Theorem to conclude that there must be some c in $[0, 1]$ such that

$$f(c) = 0 \qquad f \text{ has a zero in the closed interval } [0, 1].$$

as shown in Figure 1.37.

Try It | Exploration A

The **bisection method** for approximating the real zeros of a continuous function is similar to the method used in Example 8. If you know that a zero exists in the closed interval $[a, b]$, the zero must lie in the interval $[a, (a + b)/2]$ or $[(a + b)/2, b]$. From the sign of $f([a + b]/2)$, you can determine which interval contains the zero. By repeatedly bisecting the interval, you can "close in" on the zero of the function.

TECHNOLOGY You can also use the *zoom* feature of a graphing utility to approximate the real zeros of a continuous function. By repeatedly zooming in on the point where the graph crosses the x-axis, and adjusting the x-axis scale, you can approximate the zero of the function to any desired accuracy. The zero of $x^3 + 2x - 1$ is approximately 0.453, as shown in Figure 1.38.

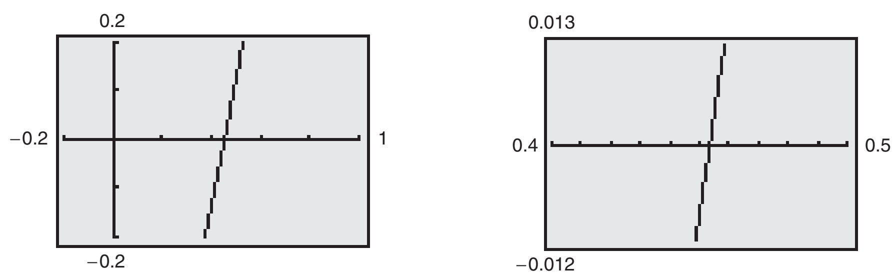


Figure 1.38 Zooming in on the zero of $f(x) = x^3 + 2x - 1$

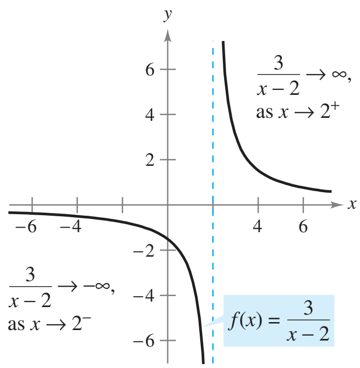
Section 1.5 Infinite Limits

- **Determine infinite limits from the left and from the right.**
- **Find and sketch the vertical asymptotes of the graph of a function.**

Infinite Limits

Let f be the function given by

$$f(x) = \frac{3}{x-2}.$$



$f(x)$ increases and decreases without bound as x approaches 2.
Figure 1.39

From Figure 1.39 and the table, you can see that $f(x)$ *decreases without bound* as x approaches 2 from the left, and $f(x)$ *increases without bound* as x approaches 2 from the right. This behavior is denoted as

$$\lim_{x\to 2^-} \frac{3}{x-2} = -\infty \qquad f(x) \text{ decreases without bound as } x \text{ approaches 2 from the left.}$$

and

$$\lim_{x\to 2^+} \frac{3}{x-2} = \infty \qquad f(x) \text{ increases without bound as } x \text{ approaches 2 from the right.}$$

x approaches 2 from the left. → ← x approaches 2 from the right.

x	1.5	1.9	1.99	1.999	2	2.001	2.01	2.1	2.5
$f(x)$	−6	−30	−300	−3000	?	3000	300	30	6

$f(x)$ decreases without bound. → ← $f(x)$ increases without bound.

A limit in which $f(x)$ increases or decreases without bound as x approaches c is called an **infinite limit.**

Definition of Infinite Limits

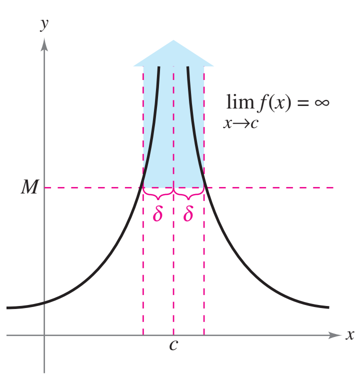
Let f be a function that is defined at every real number in some open interval containing c (except possibly at c itself). The statement

$$\lim_{x\to c} f(x) = \infty$$

means that for each $M > 0$ there exists a $\delta > 0$ such that $f(x) > M$ whenever $0 < |x - c| < \delta$ (see Figure 1.40). Similarly, the statement

$$\lim_{x\to c} f(x) = -\infty$$

means that for each $N < 0$ there exists a $\delta > 0$ such that $f(x) < N$ whenever $0 < |x - c| < \delta$. To define the **infinite limit from the left,** replace $0 < |x - c| < \delta$ by $c - \delta < x < c$. To define the **infinite limit from the right,** replace $0 < |x - c| < \delta$ by $c < x < c + \delta$.



Infinite limits
Figure 1.40

Video

Be sure you see that the equal sign in the statement $\lim f(x) = \infty$ does not mean that the limit exists! On the contrary, it tells you how the limit *fails to exist* by denoting the unbounded behavior of $f(x)$ as x approaches c.

EXPLORATION

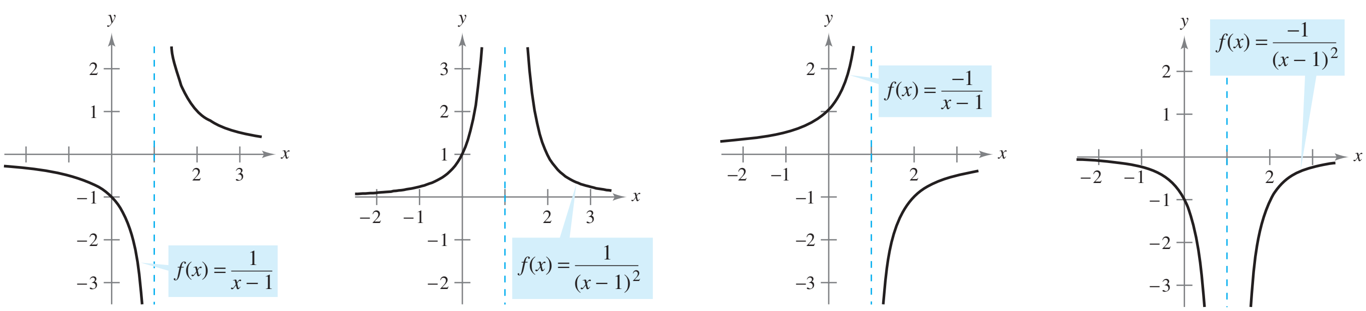
Use a graphing utility to graph each function. For each function, analytically find the single real number c that is not in the domain. Then graphically find the limit of $f(x)$ as x approaches c from the left and from the right.

a. $f(x) = \dfrac{3}{x - 4}$ **b.** $f(x) = \dfrac{1}{2 - x}$

c. $f(x) = \dfrac{2}{(x - 3)^2}$ **d.** $f(x) = \dfrac{-3}{(x + 2)^2}$

EXAMPLE 1 Determining Infinite Limits from a Graph

Use Figure 1.41 to determine the limit of each function as x approaches 1 from the left and from the right.



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

Editable Graph Editable Graph Editable Graph Editable Graph

Figure 1.41 Each graph has an asymptote at $x = 1$.

Solution

a. $\displaystyle\lim_{x\to 1^-} \frac{1}{x - 1} = -\infty$ and $\displaystyle\lim_{x\to 1^+} \frac{1}{x - 1} = \infty$

b. $\displaystyle\lim_{x\to 1} \frac{1}{(x - 1)^2} = \infty$ Limit from each side is ∞.

c. $\displaystyle\lim_{x\to 1^-} \frac{-1}{x - 1} = \infty$ and $\displaystyle\lim_{x\to 1^+} \frac{-1}{x - 1} = -\infty$

d. $\displaystyle\lim_{x\to 1} \frac{-1}{(x - 1)^2} = -\infty$ Limit from each side is $-\infty$.

Try It Exploration A

Vertical Asymptotes

If it were possible to extend the graphs in Figure 1.41 toward positive and negative infinity, you would see that each graph becomes arbitrarily close to the vertical line $x = 1$. This line is a **vertical asymptote** of the graph of f. (You will study other types of asymptotes in Sections 3.5 and 3.6.)

NOTE If the graph of a function f has a vertical asymptote at $x = c$, then f is *not continuous* at c.

Definition of Vertical Asymptote

If $f(x)$ approaches infinity (or negative infinity) as x approaches c from the right or the left, then the line $x = c$ is a **vertical asymptote** of the graph of f.

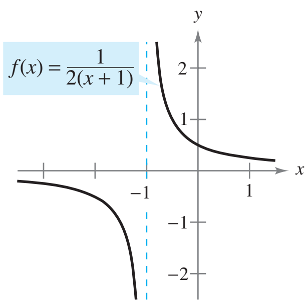
In Example 1, note that each of the functions is a *quotient* and that the vertical asymptote occurs at a number where the denominator is 0 (and the numerator is not 0). The next theorem generalizes this observation. (A proof of this theorem is given in Appendix A.)

THEOREM 1.14 Vertical Asymptotes

Let f and g be continuous on an open interval containing c. If $f(c) \neq 0$, $g(c) = 0$, and there exists an open interval containing c such that $g(x) \neq 0$ for all $x \neq c$ in the interval, then the graph of the function given by

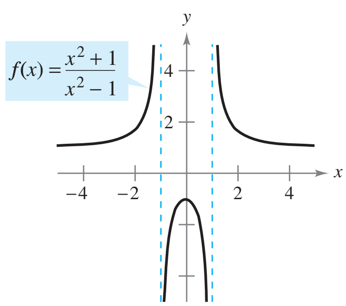
$$h(x) = \frac{f(x)}{g(x)}$$

has a vertical asymptote at $x = c$.



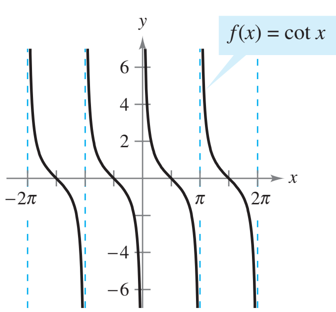
(a)

Editable Graph



(b)

Editable Graph



(c)

Editable Graph

Functions with vertical asymptotes
Figure 1.42

Video

EXAMPLE 2 Finding Vertical Asymptotes

Determine all vertical asymptotes of the graph of each function.

a. $f(x) = \dfrac{1}{2(x+1)}$ **b.** $f(x) = \dfrac{x^2+1}{x^2-1}$ **c.** $f(x) = \cot x$

Solution

a. When $x = -1$, the denominator of

$$f(x) = \frac{1}{2(x+1)}$$

is 0 and the numerator is not 0. So, by Theorem 1.14, you can conclude that $x = -1$ is a vertical asymptote, as shown in Figure 1.42(a).

b. By factoring the denominator as

$$f(x) = \frac{x^2+1}{x^2-1} = \frac{x^2+1}{(x-1)(x+1)}$$

you can see that the denominator is 0 at $x = -1$ and $x = 1$. Moreover, because the numerator is not 0 at these two points, you can apply Theorem 1.14 to conclude that the graph of f has two vertical asymptotes, as shown in Figure 1.42(b).

c. By writing the cotangent function in the form

$$f(x) = \cot x = \frac{\cos x}{\sin x}$$

you can apply Theorem 1.14 to conclude that vertical asymptotes occur at all values of x such that $\sin x = 0$ and $\cos x \neq 0$, as shown in Figure 1.42(c). So, the graph of this function has infinitely many vertical asymptotes. These asymptotes occur when $x = n\pi$, where n is an integer.

Try It | Exploration A | Exploration B | Open Exploration

Theorem 1.14 requires that the value of the numerator at $x = c$ be nonzero. If both the numerator and the denominator are 0 at $x = c$, you obtain the *indeterminate form* $0/0$, and you cannot determine the limit behavior at $x = c$ without further investigation, as illustrated in Example 3.

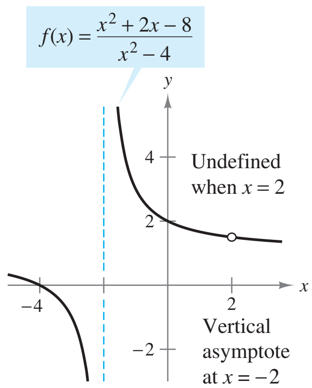
EXAMPLE 3 A Rational Function with Common Factors

Determine all vertical asymptotes of the graph of

$$f(x) = \frac{x^2 + 2x - 8}{x^2 - 4}.$$

Solution Begin by simplifying the expression, as shown.

$$\begin{aligned} f(x) &= \frac{x^2 + 2x - 8}{x^2 - 4} \\ &= \frac{(x + 4)\cancel{(x - 2)}}{(x + 2)\cancel{(x - 2)}} \\ &= \frac{x + 4}{x + 2}, \quad x \neq 2 \end{aligned}$$



$f(x)$ increases and decreases without bound as x approaches -2.
Figure 1.43

Editable Graph

At all x-values other than $x = 2$, the graph of f coincides with the graph of $g(x) = (x + 4)/(x + 2)$. So, you can apply Theorem 1.14 to g to conclude that there is a vertical asymptote at $x = -2$, as shown in Figure 1.43. From the graph, you can see that

$$\lim_{x \to -2^-} \frac{x^2 + 2x - 8}{x^2 - 4} = -\infty \quad \text{and} \quad \lim_{x \to -2^+} \frac{x^2 + 2x - 8}{x^2 - 4} = \infty.$$

Note that $x = 2$ is *not* a vertical asymptote.

Try It Exploration A Exploration B

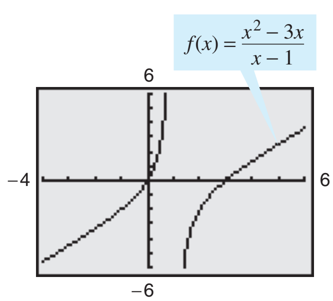
EXAMPLE 4 Determining Infinite Limits

Find each limit.

$$\lim_{x \to 1^-} \frac{x^2 - 3x}{x - 1} \quad \text{and} \quad \lim_{x \to 1^+} \frac{x^2 - 3x}{x - 1}$$

Solution Because the denominator is 0 when $x = 1$ (and the numerator is not zero), you know that the graph of

$$f(x) = \frac{x^2 - 3x}{x - 1}$$



f has a vertical asymptote at $x = 1$.
Figure 1.44

Editable Graph

has a vertical asymptote at $x = 1$. This means that each of the given limits is either ∞ or $-\infty$. A graphing utility can help determine the result. From the graph of f shown in Figure 1.44, you can see that the graph approaches ∞ from the left of $x = 1$ and approaches $-\infty$ from the right of $x = 1$. So, you can conclude that

$$\lim_{x \to 1^-} \frac{x^2 - 3x}{x - 1} = \infty \qquad \text{The limit from the left is infinity.}$$

and

$$\lim_{x \to 1^+} \frac{x^2 - 3x}{x - 1} = -\infty. \qquad \text{The limit from the right is negative infinity.}$$

Try It Exploration A

TECHNOLOGY PITFALL When using a graphing calculator or graphing software, be careful to interpret correctly the graph of a function with a vertical asymptote—graphing utilities often have difficulty drawing this type of graph.

THEOREM 1.15 Properties of Infinite Limits

Let c and L be real numbers and let f and g be functions such that

$$\lim_{x \to c} f(x) = \infty \quad \text{and} \quad \lim_{x \to c} g(x) = L.$$

1. Sum or difference: $\lim_{x \to c} [f(x) \pm g(x)] = \infty$

2. Product: $\lim_{x \to c} [f(x)g(x)] = \infty, \quad L > 0$

$\lim_{x \to c} [f(x)g(x)] = -\infty, \quad L < 0$

3. Quotient: $\lim_{x \to c} \dfrac{g(x)}{f(x)} = 0$

Similar properties hold for one-sided limits and for functions for which the limit of $f(x)$ as x approaches c is $-\infty$.

Proof To show that the limit of $f(x) + g(x)$ is infinite, choose $M > 0$. You then need to find $\delta > 0$ such that

$$[f(x) + g(x)] > M$$

whenever $0 < |x - c| < \delta$. For simplicity's sake, you can assume L is positive. Let $M_1 = M + 1$. Because the limit of $f(x)$ is infinite, there exists δ_1 such that $f(x) > M_1$ whenever $0 < |x - c| < \delta_1$. Also, because the limit of $g(x)$ is L, there exists δ_2 such that $|g(x) - L| < 1$ whenever $0 < |x - c| < \delta_2$. By letting δ be the smaller of δ_1 and δ_2, you can conclude that $0 < |x - c| < \delta$ implies $f(x) > M + 1$ and $|g(x) - L| < 1$. The second of these two inequalities implies that $g(x) > L - 1$, and, adding this to the first inequality, you can write

$$f(x) + g(x) > (M + 1) + (L - 1) = M + L > M.$$

So, you can conclude that

$$\lim_{x \to c} [f(x) + g(x)] = \infty.$$

The proofs of the remaining properties are left as exercises (see Exercise 72).

EXAMPLE 5 Determining Limits

a. Because $\lim_{x \to 0} 1 = 1$ and $\lim_{x \to 0} \dfrac{1}{x^2} = \infty$, you can write

$$\lim_{x \to 0} \left(1 + \frac{1}{x^2}\right) = \infty.$$ Property 1, Theorem 1.15

b. Because $\lim_{x \to 1^-} (x^2 + 1) = 2$ and $\lim_{x \to 1^-} (\cot \pi x) = -\infty$, you can write

$$\lim_{x \to 1^-} \frac{x^2 + 1}{\cot \pi x} = 0.$$ Property 3, Theorem 1.15

c. Because $\lim_{x \to 0^+} 3 = 3$ and $\lim_{x \to 0^+} \cot x = \infty$, you can write

$$\lim_{x \to 0^+} 3 \cot x = \infty.$$ Property 2, Theorem 1.15

Try It

Exploration A

Section 2.1 The Derivative and the Tangent Line Problem

- **Find the slope of the tangent line to a curve at a point.**
- **Use the limit definition to find the derivative of a function.**
- **Understand the relationship between differentiability and continuity.**

ISAAC NEWTON (1642–1727)
In addition to his work in calculus, Newton made revolutionary contributions to physics, including the Law of Universal Gravitation and his three laws of motion.

MathBio

The Tangent Line Problem

Calculus grew out of four major problems that European mathematicians were working on during the seventeenth century.

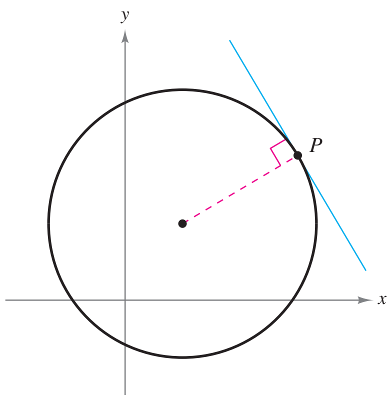
1. The tangent line problem (Section 1.1 and this section)
2. The velocity and acceleration problem (Sections 2.2 and 2.3)
3. The minimum and maximum problem (Section 3.1)
4. The area problem (Sections 1.1 and 4.2)

Each problem involves the notion of a limit, and calculus can be introduced with any of the four problems.

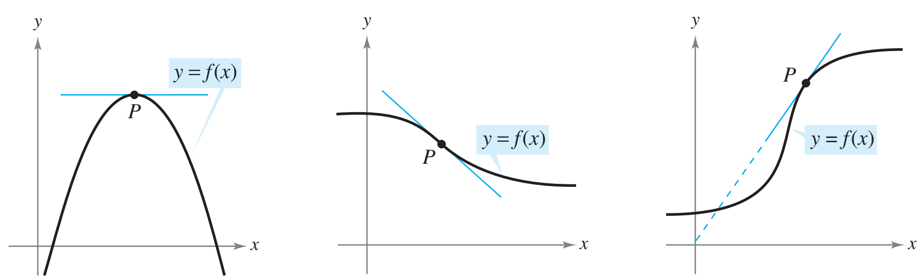
A brief introduction to the tangent line problem is given in Section 1.1. Although partial solutions to this problem were given by Pierre de Fermat (1601–1665), René Descartes (1596–1650), Christian Huygens (1629–1695), and Isaac Barrow (1630–1677), credit for the first general solution is usually given to Isaac Newton (1642–1727) and Gottfried Leibniz (1646–1716). Newton's work on this problem stemmed from his interest in optics and light refraction.

What does it mean to say that a line is tangent to a curve at a point? For a circle, the tangent line at a point P is the line that is perpendicular to the radial line at point P, as shown in Figure 2.1.

For a general curve, however, the problem is more difficult. For example, how would you define the tangent lines shown in Figure 2.2? You might say that a line is tangent to a curve at a point P if it touches, but does not cross, the curve at point P. This definition would work for the first curve shown in Figure 2.2, but not for the second. *Or* you might say that a line is tangent to a curve if the line touches or intersects the curve at exactly one point. This definition would work for a circle but not for more general curves, as the third curve in Figure 2.2 shows.



Tangent line to a circle
Figure 2.1



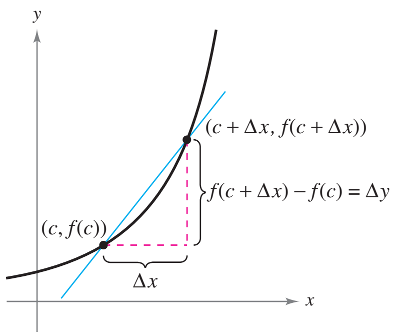
Tangent line to a curve at a point
Figure 2.2

FOR FURTHER INFORMATION For more information on the crediting of mathematical discoveries to the first "discoverer," see the article "Mathematical Firsts—Who Done It?" by Richard H. Williams and Roy D. Mazzagatti in *Mathematics Teacher*.

MathArticle

EXPLORATION

Identifying a Tangent Line Use a graphing utility to graph the function $f(x) = 2x^3 - 4x^2 + 3x - 5$. On the same screen, graph $y = x - 5$, $y = 2x - 5$, and $y = 3x - 5$. Which of these lines, if any, appears to be tangent to the graph of f at the point $(0, -5)$? Explain your reasoning.



The secant line through $(c, f(c))$ and $(c + \Delta x, f(c + \Delta x))$
Figure 2.3

Essentially, the problem of finding the tangent line at a point P boils down to the problem of finding the *slope* of the tangent line at point P. You can approximate this slope using a **secant line*** through the point of tangency and a second point on the curve, as shown in Figure 2.3. If $(c, f(c))$ is the point of tangency and $(c + \Delta x, f(c + \Delta x))$ is a second point on the graph of f, the slope of the secant line through the two points is given by substitution into the slope formula

$$m = \frac{y_2 - y_1}{x_2 - x_1}$$

$$m_{\text{sec}} = \frac{f(c + \Delta x) - f(c)}{(c + \Delta x) - c} \qquad \frac{\text{Change in } y}{\text{Change in } x}$$

$$\boxed{m_{\text{sec}} = \frac{f(c + \Delta x) - f(c)}{\Delta x}.} \qquad \text{Slope of secant line}$$

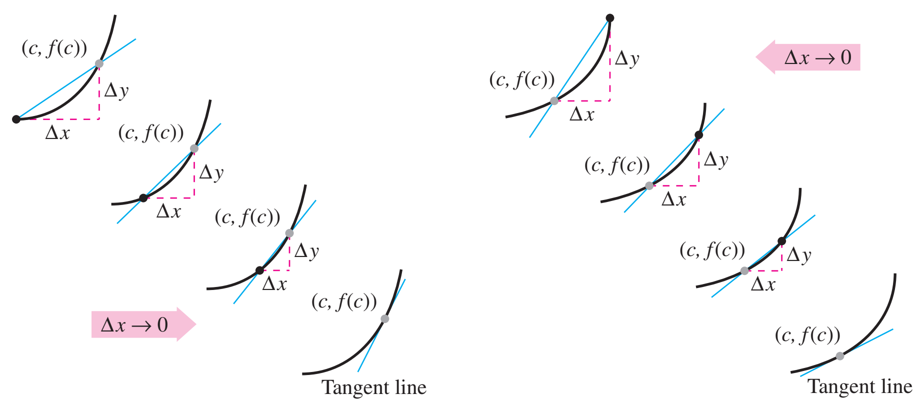
The right-hand side of this equation is a **difference quotient.** The denominator Δx is the **change in x,** and the numerator $\Delta y = f(c + \Delta x) - f(c)$ is the **change in y.**

The beauty of this procedure is that you can obtain more and more accurate approximations of the slope of the tangent line by choosing points closer and closer to the point of tangency, as shown in Figure 2.4.

THE TANGENT LINE PROBLEM

In 1637, mathematician René Descartes stated this about the tangent line problem:

"And I dare say that this is not only the most useful and general problem in geometry that I know, but even that I ever desire to know."



Tangent line approximations
Figure 2.4

To view a sequence of secant lines approaching a tangent line, select the Animation button.

Animation

Definition of Tangent Line with Slope m

If f is defined on an open interval containing c, and if the limit

$$\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \to 0} \frac{f(c + \Delta x) - f(c)}{\Delta x} = m$$

exists, then the line passing through $(c, f(c))$ with slope m is the **tangent line** to the graph of f at the point $(c, f(c))$.

Video Video

The slope of the tangent line to the graph of f at the point $(c, f(c))$ is also called the **slope of the graph of f at $x = c$.**

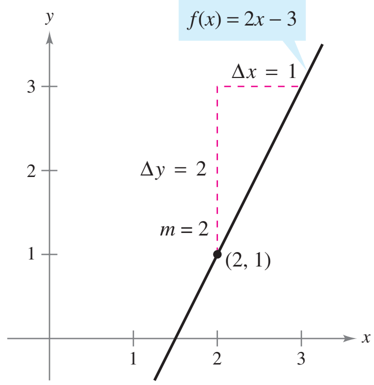
* *This use of the word* secant *comes from the Latin* secare, *meaning to cut, and is not a reference to the trigonometric function of the same name.*

EXAMPLE 1 The Slope of the Graph of a Linear Function

Find the slope of the graph of

$$f(x) = 2x - 3$$

at the point (2, 1).



The slope of f at (2, 1) is $m = 2$.
Figure 2.5

Editable Graph

Solution To find the slope of the graph of f when $c = 2$, you can apply the definition of the slope of a tangent line, as shown.

$$\begin{aligned}
\lim_{\Delta x \to 0} \frac{f(2 + \Delta x) - f(2)}{\Delta x} &= \lim_{\Delta x \to 0} \frac{[2(2 + \Delta x) - 3] - [2(2) - 3]}{\Delta x} \\
&= \lim_{\Delta x \to 0} \frac{4 + 2\Delta x - 3 - 4 + 3}{\Delta x} \\
&= \lim_{\Delta x \to 0} \frac{2\cancel{\Delta x}}{\cancel{\Delta x}} \\
&= \lim_{\Delta x \to 0} 2 \\
&= 2
\end{aligned}$$

The slope of f at $(c, f(c)) = (2, 1)$ is $m = 2$, as shown in Figure 2.5.

NOTE In Example 1, the limit definition of the slope of f agrees with the definition of the slope of a line as discussed in Section P.2.

Try It Exploration A

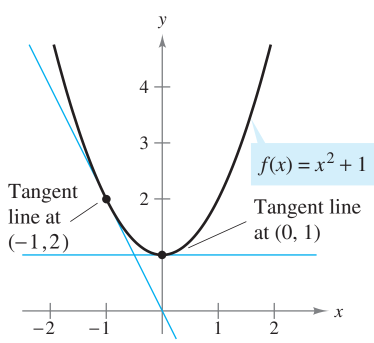
The graph of a linear function has the same slope at any point. This is not true of nonlinear functions, as shown in the following example.

EXAMPLE 2 Tangent Lines to the Graph of a Nonlinear Function

Find the slopes of the tangent lines to the graph of

$$f(x) = x^2 + 1$$

at the points (0, 1) and $(-1, 2)$, as shown in Figure 2.6.



The slope of f at any point $(c, f(c))$ is $m = 2c$.
Figure 2.6

Editable Graph

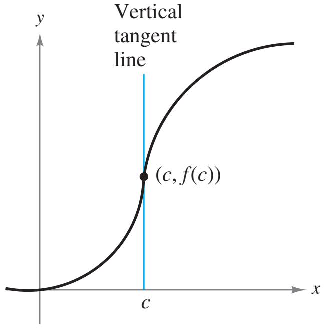
Solution Let $(c, f(c))$ represent an arbitrary point on the graph of f. Then the slope of the tangent line at $(c, f(c))$ is given by

$$\begin{aligned}
\lim_{\Delta x \to 0} \frac{f(c + \Delta x) - f(c)}{\Delta x} &= \lim_{\Delta x \to 0} \frac{[(c + \Delta x)^2 + 1] - (c^2 + 1)}{\Delta x} \\
&= \lim_{\Delta x \to 0} \frac{c^2 + 2c(\Delta x) + (\Delta x)^2 + 1 - c^2 - 1}{\Delta x} \\
&= \lim_{\Delta x \to 0} \frac{2c(\Delta x) + (\Delta x)^2}{\Delta x} \\
&= \lim_{\Delta x \to 0} (2c + \Delta x) \\
&= 2c.
\end{aligned}$$

So, the slope at *any* point $(c, f(c))$ on the graph of f is $m = 2c$. At the point (0, 1), the slope is $m = 2(0) = 0$, and at $(-1, 2)$, the slope is $m = 2(-1) = -2$.

NOTE In Example 2, note that c is held constant in the limit process (as $\Delta x \to 0$).

Try It Exploration A



The graph of f has a vertical tangent line at $(c, f(c))$.
Figure 2.7

The definition of a tangent line to a curve does not cover the possibility of a vertical tangent line. For vertical tangent lines, you can use the following definition. If f is continuous at c and

$$\lim_{\Delta x \to 0} \frac{f(c + \Delta x) - f(c)}{\Delta x} = \infty \quad \text{or} \quad \lim_{\Delta x \to 0} \frac{f(c + \Delta x) - f(c)}{\Delta x} = -\infty$$

the vertical line $x = c$ passing through $(c, f(c))$ is a **vertical tangent line** to the graph of f. For example, the function shown in Figure 2.7 has a vertical tangent line at $(c, f(c))$. If the domain of f is the closed interval $[a, b]$, you can extend the definition of a vertical tangent line to include the endpoints by considering continuity and limits from the right (for $x = a$) and from the left (for $x = b$).

The Derivative of a Function

You have now arrived at a crucial point in the study of calculus. The limit used to define the slope of a tangent line is also used to define one of the two fundamental operations of calculus—**differentiation.**

Definition of the Derivative of a Function

The **derivative** of f at x is given by

$$f'(x) = \lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}$$

provided the limit exists. For all x for which this limit exists, f' is a function of x.

Video

Be sure you see that the derivative of a function of x is also a function of x. This "new" function gives the slope of the tangent line to the graph of f at the point $(x, f(x))$, provided that the graph has a tangent line at this point.

The process of finding the derivative of a function is called **differentiation.** A function is **differentiable** at x if its derivative exists at x and is **differentiable on an open interval (a, b)** if it is differentiable at every point in the interval.

In addition to $f'(x)$, which is read as "f prime of x," other notations are used to denote the derivative of $y = f(x)$. The most common are

$$f'(x), \quad \frac{dy}{dx}, \quad y', \quad \frac{d}{dx}[f(x)], \quad D_x[y].$$

Notation for derivatives

The notation dy/dx is read as "the derivative of y *with respect to* x" or simply "$dy - dx$". Using limit notation, you can write

$$\begin{aligned} \frac{dy}{dx} &= \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} \\ &= \lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x} \\ &= f'(x). \end{aligned}$$

History

EXAMPLE 3 Finding the Derivative by the Limit Process

Find the derivative of $f(x) = x^3 + 2x$.

Solution

$$\begin{aligned}
f'(x) &= \lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x} && \text{Definition of derivative} \\
&= \lim_{\Delta x \to 0} \frac{(x + \Delta x)^3 + 2(x + \Delta x) - (x^3 + 2x)}{\Delta x} \\
&= \lim_{\Delta x \to 0} \frac{x^3 + 3x^2\Delta x + 3x(\Delta x)^2 + (\Delta x)^3 + 2x + 2\Delta x - x^3 - 2x}{\Delta x} \\
&= \lim_{\Delta x \to 0} \frac{3x^2\Delta x + 3x(\Delta x)^2 + (\Delta x)^3 + 2\Delta x}{\Delta x} \\
&= \lim_{\Delta x \to 0} \frac{\cancel{\Delta x}[3x^2 + 3x\Delta x + (\Delta x)^2 + 2]}{\cancel{\Delta x}} \\
&= \lim_{\Delta x \to 0} [3x^2 + 3x\Delta x + (\Delta x)^2 + 2] \\
&= 3x^2 + 2
\end{aligned}$$

Try It | Exploration A | Exploration B

Exploration C | Open Exploration

STUDY TIP When using the definition to find a derivative of a function, the key is to rewrite the difference quotient so that Δx does not occur as a factor of the denominator.

The editable graph feature below allows you to edit the graph of a function and its derivative.

Editable Graph

Remember that the derivative of a function f is itself a function, which can be used to find the slope of the tangent line at the point $(x, f(x))$ on the graph of f.

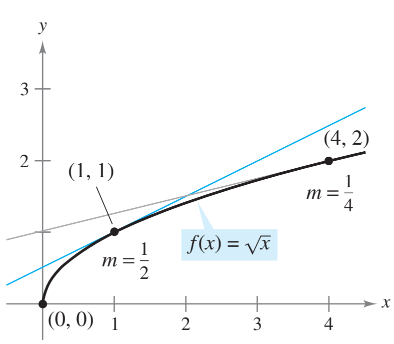
EXAMPLE 4 Using the Derivative to Find the Slope at a Point

Find $f'(x)$ for $f(x) = \sqrt{x}$. Then find the slope of the graph of f at the points $(1, 1)$ and $(4, 2)$. Discuss the behavior of f at $(0, 0)$.

Solution Use the procedure for rationalizing numerators, as discussed in Section 1.3.

$$\begin{aligned}
f'(x) &= \lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x} && \text{Definition of derivative} \\
&= \lim_{\Delta x \to 0} \frac{\sqrt{x + \Delta x} - \sqrt{x}}{\Delta x} \\
&= \lim_{\Delta x \to 0} \left(\frac{\sqrt{x + \Delta x} - \sqrt{x}}{\Delta x}\right)\left(\frac{\sqrt{x + \Delta x} + \sqrt{x}}{\sqrt{x + \Delta x} + \sqrt{x}}\right) \\
&= \lim_{\Delta x \to 0} \frac{(x + \Delta x) - x}{\Delta x\left(\sqrt{x + \Delta x} + \sqrt{x}\right)} \\
&= \lim_{\Delta x \to 0} \frac{\cancel{\Delta x}}{\cancel{\Delta x}\left(\sqrt{x + \Delta x} + \sqrt{x}\right)} \\
&= \lim_{\Delta x \to 0} \frac{1}{\sqrt{x + \Delta x} + \sqrt{x}} \\
&= \frac{1}{2\sqrt{x}}, \quad x > 0
\end{aligned}$$

At the point $(1, 1)$, the slope is $f'(1) = \frac{1}{2}$. At the point $(4, 2)$, the slope is $f'(4) = \frac{1}{4}$. See Figure 2.8. At the point $(0, 0)$, the slope is undefined. Moreover, the graph of f has a vertical tangent line at $(0, 0)$.



The slope of f at $(x, f(x))$, $x > 0$, is $m = 1/(2\sqrt{x})$.
Figure 2.8

Editable Graph

Try It | Exploration A | Exploration B | Exploration C

In many applications, it is convenient to use a variable other than x as the independent variable, as shown in Example 5.

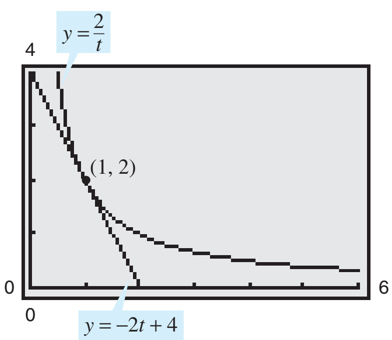
EXAMPLE 5 Finding the Derivative of a Function

Find the derivative with respect to t for the function $y = 2/t$.

Solution Considering $y = f(t)$, you obtain

$$\begin{aligned}
\frac{dy}{dt} &= \lim_{\Delta t \to 0} \frac{f(t + \Delta t) - f(t)}{\Delta t} && \text{Definition of derivative} \\
&= \lim_{\Delta t \to 0} \frac{\dfrac{2}{t + \Delta t} - \dfrac{2}{t}}{\Delta t} && f(t + \Delta t) = 2/(t + \Delta t) \text{ and } f(t) = 2/t \\
&= \lim_{\Delta t \to 0} \frac{\dfrac{2t - 2(t + \Delta t)}{t(t + \Delta t)}}{\Delta t} && \text{Combine fractions in numerator.} \\
&= \lim_{\Delta t \to 0} \frac{-2\cancel{\Delta t}}{\cancel{\Delta t}(t)(t + \Delta t)} && \text{Divide out common factor of } \Delta t. \\
&= \lim_{\Delta t \to 0} \frac{-2}{t(t + \Delta t)} && \text{Simplify.} \\
&= -\frac{2}{t^2}. && \text{Evaluate limit as } \Delta t \to 0.
\end{aligned}$$

Try It | Exploration A | Open Exploration



At the point $(1, 2)$ the line $y = -2t + 4$ is tangent to the graph of $y = 2/t$.
Figure 2.9

The editable graph feature below allows you to edit the graph of a function and its derivative.

Editable Graph

TECHNOLOGY A graphing utility can be used to reinforce the result given in Example 5. For instance, using the formula $dy/dt = -2/t^2$, you know that the slope of the graph of $y = 2/t$ at the point $(1, 2)$ is $m = -2$. This implies that an equation of the tangent line to the graph at $(1, 2)$ is

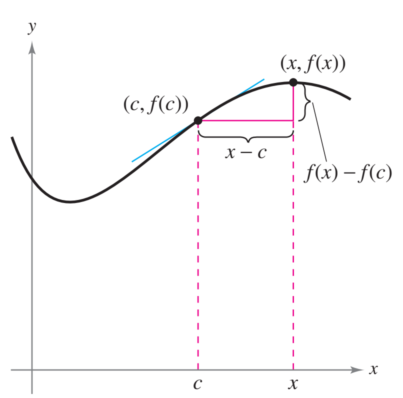
$$y - 2 = -2(t - 1) \quad \text{or} \quad y = -2t + 4$$

as shown in Figure 2.9.

Differentiability and Continuity

The following alternative limit form of the derivative is useful in investigating the relationship between differentiability and continuity. The derivative of f at c is

$$\boxed{f'(c) = \lim_{x \to c} \frac{f(x) - f(c)}{x - c}} \qquad \text{Alternative form of derivative}$$

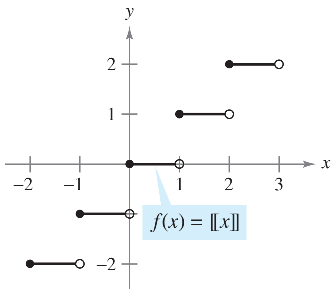


As x approaches c, the secant line approaches the tangent line.
Figure 2.10

provided this limit exists (see Figure 2.10). (A proof of the equivalence of this form is given in Appendix A.) Note that the existence of the limit in this alternative form requires that the one-sided limits

$$\lim_{x \to c^-} \frac{f(x) - f(c)}{x - c} \quad \text{and} \quad \lim_{x \to c^+} \frac{f(x) - f(c)}{x - c}$$

exist and are equal. These one-sided limits are called the **derivatives from the left and from the right,** respectively. It follows that f is **differentiable on the closed interval** $[a, b]$ if it is differentiable on (a, b) and if the derivative from the right at a and the derivative from the left at b both exist.



The greatest integer function is not differentiable at $x = 0$, because it is not continuous at $x = 0$.
Figure 2.11

If a function is not continuous at $x = c$, it is also not differentiable at $x = c$. For instance, the greatest integer function

$$f(x) = [\![x]\!]$$

is not continuous at $x = 0$, and so it is not differentiable at $x = 0$ (see Figure 2.11). You can verify this by observing that

$$\lim_{x \to 0^-} \frac{f(x) - f(0)}{x - 0} = \lim_{x \to 0^-} \frac{[\![x]\!] - 0}{x} = \infty \qquad \text{Derivative from the left}$$

and

$$\lim_{x \to 0^+} \frac{f(x) - f(0)}{x - 0} = \lim_{x \to 0^+} \frac{[\![x]\!] - 0}{x} = 0. \qquad \text{Derivative from the right}$$

Although it is true that differentiability implies continuity (as shown in Theorem 2.1 on the next page), the converse is not true. That is, it is possible for a function to be continuous at $x = c$ and *not* differentiable at $x = c$. Examples 6 and 7 illustrate this possibility.

EXAMPLE 6 A Graph with a Sharp Turn

The function

$$f(x) = |x - 2|$$

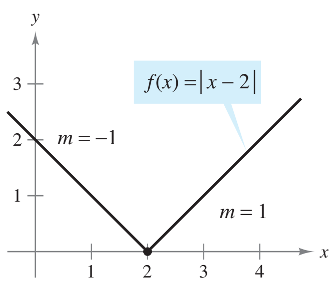
shown in Figure 2.12 is continuous at $x = 2$. But, the one-sided limits

$$\lim_{x \to 2^-} \frac{f(x) - f(2)}{x - 2} = \lim_{x \to 2^-} \frac{|x - 2| - 0}{x - 2} = -1 \qquad \text{Derivative from the left}$$

and

$$\lim_{x \to 2^+} \frac{f(x) - f(2)}{x - 2} = \lim_{x \to 2^+} \frac{|x - 2| - 0}{x - 2} = 1 \qquad \text{Derivative from the right}$$

are not equal. So, f is not differentiable at $x = 2$ and the graph of f does not have a tangent line at the point $(2, 0)$.



f is not differentiable at $x = 2$, because the derivatives from the left and from the right are not equal.
Figure 2.12

Editable Graph

Try It | Exploration A | Open Exploration

EXAMPLE 7 A Graph with a Vertical Tangent Line

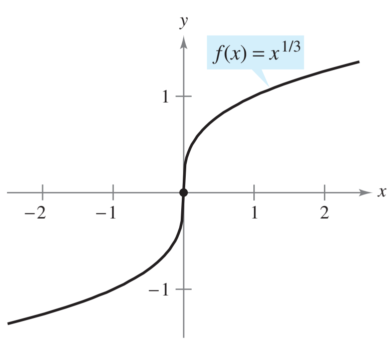
The function

$$f(x) = x^{1/3}$$

is continuous at $x = 0$, as shown in Figure 2.13. But, because the limit

$$\lim_{x \to 0} \frac{f(x) - f(0)}{x - 0} = \lim_{x \to 0} \frac{x^{1/3} - 0}{x}$$
$$= \lim_{x \to 0} \frac{1}{x^{2/3}}$$
$$= \infty$$

is infinite, you can conclude that the tangent line is vertical at $x = 0$. So, f is not differentiable at $x = 0$.



f is not differentiable at $x = 0$, because f has a vertical tangent at $x = 0$.
Figure 2.13

Editable Graph

Try It | Exploration A | Exploration B | Exploration C

From Examples 6 and 7, you can see that a function is not differentiable at a point at which its graph has a sharp turn *or* a vertical tangent.

TECHNOLOGY Some graphing utilities, such as *Derive*, *Maple*, *Mathcad*, *Mathematica*, and the *TI-89*, perform symbolic differentiation. Others perform *numerical differentiation* by finding values of derivatives using the formula

$$f'(x) \approx \frac{f(x + \Delta x) - f(x - \Delta x)}{2\Delta x}$$

where Δx is a small number such as 0.001. Can you see any problems with this definition? For instance, using this definition, what is the value of the derivative of $f(x) = |x|$ when $x = 0$?

THEOREM 2.1 Differentiability Implies Continuity

If f is differentiable at $x = c$, then f is continuous at $x = c$.

Proof You can prove that f is continuous at $x = c$ by showing that $f(x)$ approaches $f(c)$ as $x \to c$. To do this, use the differentiability of f at $x = c$ and consider the following limit.

$$\begin{aligned}
\lim_{x\to c} [f(x) - f(c)] &= \lim_{x\to c} \left[(x - c)\left(\frac{f(x) - f(c)}{x - c}\right)\right] \\
&= \left[\lim_{x\to c} (x - c)\right]\left[\lim_{x\to c} \frac{f(x) - f(c)}{x - c}\right] \\
&= (0)[f'(c)] \\
&= 0
\end{aligned}$$

Because the difference $f(x) - f(c)$ approaches zero as $x \to c$, you can conclude that $\lim_{x\to c} f(x) = f(c)$. So, f is continuous at $x = c$.

The following statements summarize the relationship between continuity and differentiability.

1. If a function is differentiable at $x = c$, then it is continuous at $x = c$. So, differentiability implies continuity.
2. It is possible for a function to be continuous at $x = c$ and not be differentiable at $x = c$. So, continuity does not imply differentiability.

Section 2.2 Basic Differentiation Rules and Rates of Change

Video Video
Video Video

- Find the derivative of a function using the Constant Rule.
- Find the derivative of a function using the Power Rule.
- Find the derivative of a function using the Constant Multiple Rule.
- Find the derivative of a function using the Sum and Difference Rules.
- Find the derivatives of the sine function and of the cosine function.
- Use derivatives to find rates of change.

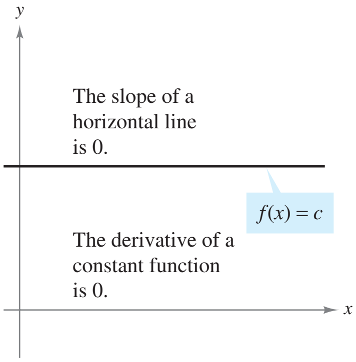
The Constant Rule

In Section 2.1 you used the limit definition to find derivatives. In this and the next two sections you will be introduced to several "differentiation rules" that allow you to find derivatives without the *direct* use of the limit definition.

THEOREM 2.2 The Constant Rule

The derivative of a constant function is 0. That is, if c is a real number, then

$$\frac{d}{dx}[c] = 0.$$



The Constant Rule
Figure 2.14

NOTE In Figure 2.14, note that the Constant Rule is equivalent to saying that the slope of a horizontal line is 0. This demonstrates the relationship between slope and derivative.

Proof Let $f(x) = c$. Then, by the limit definition of the derivative,

$$\begin{aligned}\frac{d}{dx}[c] &= f'(x)\\ &= \lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}\\ &= \lim_{\Delta x \to 0} \frac{c - c}{\Delta x}\\ &= \lim_{\Delta x \to 0} 0\\ &= 0.\end{aligned}$$

EXAMPLE 1 Using the Constant Rule

Function	*Derivative*
a. $y = 7$	$\dfrac{dy}{dx} = 0$
b. $f(x) = 0$	$f'(x) = 0$
c. $s(t) = -3$	$s'(t) = 0$
d. $y = k\pi^2$, k is constant	$y' = 0$

Try It Exploration A

The editable graph feature below allows you to edit the graph of a function.

a. Editable Graph **b.** Editable Graph

c. Editable Graph **d.** Editable Graph

EXPLORATION

Writing a Conjecture Use the definition of the derivative given in Section 2.1 to find the derivative of each function. What patterns do you see? Use your results to write a conjecture about the derivative of $f(x) = x^n$.

a. $f(x) = x^1$ **b.** $f(x) = x^2$ **c.** $f(x) = x^3$
d. $f(x) = x^4$ **e.** $f(x) = x^{1/2}$ **f.** $f(x) = x^{-1}$

The Power Rule

Before proving the next rule, it is important to review the procedure for expanding a binomial.

$$(x + \Delta x)^2 = x^2 + 2x\Delta x + (\Delta x)^2$$
$$(x + \Delta x)^3 = x^3 + 3x^2\Delta x + 3x(\Delta x)^2 + (\Delta x)^3$$

The general binomial expansion for a positive integer n is

$$(x + \Delta x)^n = x^n + nx^{n-1}(\Delta x) + \underbrace{\frac{n(n-1)x^{n-2}}{2}(\Delta x)^2 + \cdots + (\Delta x)^n}_{(\Delta x)^2 \text{ is a factor of these terms.}}.$$

This binomial expansion is used in proving a special case of the Power Rule.

THEOREM 2.3 The Power Rule

If n is a rational number, then the function $f(x) = x^n$ is differentiable and

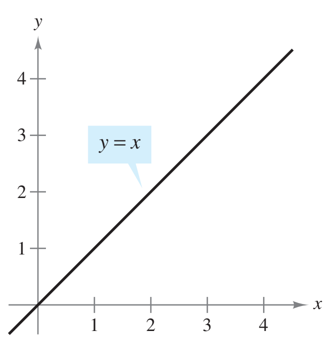
$$\frac{d}{dx}[x^n] = nx^{n-1}.$$

For f to be differentiable at $x = 0$, n must be a number such that x^{n-1} is defined on an interval containing 0.

Proof If n is a positive integer greater than 1, then the binomial expansion produces

$$\begin{aligned}
\frac{d}{dx}[x^n] &= \lim_{\Delta x \to 0} \frac{(x + \Delta x)^n - x^n}{\Delta x} \\
&= \lim_{\Delta x \to 0} \frac{x^n + nx^{n-1}(\Delta x) + \frac{n(n-1)x^{n-2}}{2}(\Delta x)^2 + \cdots + (\Delta x)^n - x^n}{\Delta x} \\
&= \lim_{\Delta x \to 0} \left[nx^{n-1} + \frac{n(n-1)x^{n-2}}{2}(\Delta x) + \cdots + (\Delta x)^{n-1}\right] \\
&= nx^{n-1} + 0 + \cdots + 0 \\
&= nx^{n-1}.
\end{aligned}$$

This proves the case for which n is a positive integer greater than 1. You will prove the case for $n = 1$. Example 7 in Section 2.3 proves the case for which n is a negative integer. In Exercise 75 in Section 2.5 you are asked to prove the case for which n is rational. (In Section 5.5, the Power Rule will be extended to cover irrational values of n.)



The slope of the line $y = x$ is 1.
Figure 2.15

When using the Power Rule, the case for which $n = 1$ is best thought of as a separate differentiation rule. That is,

$$\boxed{\frac{d}{dx}[x] = 1.}$$

Power Rule when $n = 1$

This rule is consistent with the fact that the slope of the line $y = x$ is 1, as shown in Figure 2.15.

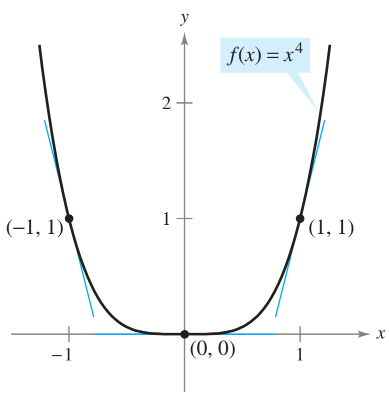
EXAMPLE 2 Using the Power Rule

Function	*Derivative*
a. $f(x) = x^3$	$f'(x) = 3x^2$
b. $g(x) = \sqrt[3]{x}$	$g'(x) = \dfrac{d}{dx}[x^{1/3}] = \dfrac{1}{3}x^{-2/3} = \dfrac{1}{3x^{2/3}}$
c. $y = \dfrac{1}{x^2}$	$\dfrac{dy}{dx} = \dfrac{d}{dx}[x^{-2}] = (-2)x^{-3} = -\dfrac{2}{x^3}$

Try It | Exploration A

In Example 2(c), note that *before* differentiating, $1/x^2$ was rewritten as x^{-2}. Rewriting is the first step in *many* differentiation problems.

Given:		Rewrite:		Differentiate:		Simplify:
$y = \dfrac{1}{x^2}$	⇨	$y = x^{-2}$	⇨	$\dfrac{dy}{dx} = (-2)x^{-3}$	⇨	$\dfrac{dy}{dx} = -\dfrac{2}{x^3}$



Note that the slope of the graph is negative at the point $(-1, 1)$, the slope is zero at the point $(0, 0)$, and the slope is positive at the point $(1, 1)$.
Figure 2.16

Editable Graph

EXAMPLE 3 Finding the Slope of a Graph

Find the slope of the graph of $f(x) = x^4$ when

a. $x = -1$ **b.** $x = 0$ **c.** $x = 1$.

Solution The slope of a graph at a point is the value of the derivative at that point. The derivative of f is $f'(x) = 4x^3$.

a. When $x = -1$, the slope is $f'(-1) = 4(-1)^3 = -4$. Slope is negative.

b. When $x = 0$, the slope is $f'(0) = 4(0)^3 = 0$. Slope is zero.

c. When $x = 1$, the slope is $f'(1) = 4(1)^3 = 4$. Slope is positive.

See Figure 2.16.

Try It | Exploration A | Open Exploration

EXAMPLE 4 Finding an Equation of a Tangent Line

Find an equation of the tangent line to the graph of $f(x) = x^2$ when $x = -2$.

Solution To find the *point* on the graph of f, evaluate the original function at $x = -2$.

$$(-2, f(-2)) = (-2, 4) \qquad \text{Point on graph}$$

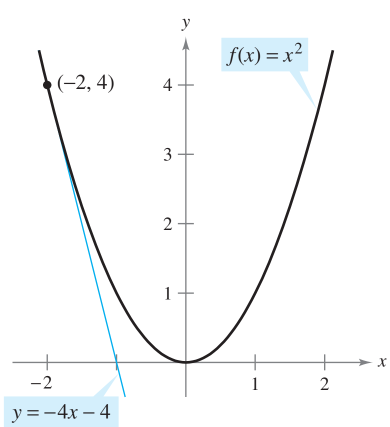
To find the *slope* of the graph when $x = -2$, evaluate the derivative, $f'(x) = 2x$, at $x = -2$.

$$m = f'(-2) = -4 \qquad \text{Slope of graph at } (-2, 4)$$

Now, using the point-slope form of the equation of a line, you can write

$$y - y_1 = m(x - x_1) \qquad \text{Point-slope form}$$
$$y - 4 = -4[x - (-2)] \qquad \text{Substitute for } y_1, m, \text{ and } x_1.$$
$$y = -4x - 4. \qquad \text{Simplify.}$$

See Figure 2.17.



The line $y = -4x - 4$ is tangent to the graph of $f(x) = x^2$ at the point $(-2, 4)$.
Figure 2.17

Editable Graph

Try It | Exploration A | Exploration B | Open Exploration

The Constant Multiple Rule

THEOREM 2.4 The Constant Multiple Rule

If f is a differentiable function and c is a real number, then cf is also differentiable and $\frac{d}{dx}[cf(x)] = cf'(x)$.

Proof

$$\begin{aligned}\frac{d}{dx}[cf(x)] &= \lim_{\Delta x \to 0} \frac{cf(x+\Delta x) - cf(x)}{\Delta x} && \text{Definition of derivative}\\ &= \lim_{\Delta x \to 0} c\left[\frac{f(x+\Delta x) - f(x)}{\Delta x}\right] \\ &= c\left[\lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x}\right] && \text{Apply Theorem 1.2.}\\ &= cf'(x)\end{aligned}$$

Informally, the Constant Multiple Rule states that constants can be factored out of the differentiation process, even if the constants appear in the denominator.

$$\frac{d}{dx}[cf(x)] = c\frac{d}{dx}[f(x)] = cf'(x)$$

$$\begin{aligned}\frac{d}{dx}\left[\frac{f(x)}{c}\right] &= \frac{d}{dx}\left[\left(\frac{1}{c}\right)f(x)\right]\\ &= \left(\frac{1}{c}\right)\frac{d}{dx}[f(x)] = \left(\frac{1}{c}\right)f'(x)\end{aligned}$$

EXAMPLE 5 Using the Constant Multiple Rule

Function	*Derivative*
a. $y = \frac{2}{x}$	$\frac{dy}{dx} = \frac{d}{dx}[2x^{-1}] = 2\frac{d}{dx}[x^{-1}] = 2(-1)x^{-2} = -\frac{2}{x^2}$
b. $f(t) = \frac{4t^2}{5}$	$f'(t) = \frac{d}{dt}\left[\frac{4}{5}t^2\right] = \frac{4}{5}\frac{d}{dt}[t^2] = \frac{4}{5}(2t) = \frac{8}{5}t$
c. $y = 2\sqrt{x}$	$\frac{dy}{dx} = \frac{d}{dx}[2x^{1/2}] = 2\left(\frac{1}{2}x^{-1/2}\right) = x^{-1/2} = \frac{1}{\sqrt{x}}$
d. $y = \frac{1}{2\sqrt[3]{x^2}}$	$\frac{dy}{dx} = \frac{d}{dx}\left[\frac{1}{2}x^{-2/3}\right] = \frac{1}{2}\left(-\frac{2}{3}\right)x^{-5/3} = -\frac{1}{3x^{5/3}}$
e. $y = -\frac{3x}{2}$	$y' = \frac{d}{dx}\left[-\frac{3}{2}x\right] = -\frac{3}{2}(1) = -\frac{3}{2}$

Try It **Exploration A**

The Constant Multiple Rule and the Power Rule can be combined into one rule. The combination rule is

$$D_x[cx^n] = cnx^{n-1}.$$

EXAMPLE 6 Using Parentheses When Differentiating

	Original Function	*Rewrite*	*Differentiate*	*Simplify*
a.	$y = \dfrac{5}{2x^3}$	$y = \dfrac{5}{2}(x^{-3})$	$y' = \dfrac{5}{2}(-3x^{-4})$	$y' = -\dfrac{15}{2x^4}$
b.	$y = \dfrac{5}{(2x)^3}$	$y = \dfrac{5}{8}(x^{-3})$	$y' = \dfrac{5}{8}(-3x^{-4})$	$y' = -\dfrac{15}{8x^4}$
c.	$y = \dfrac{7}{3x^{-2}}$	$y = \dfrac{7}{3}(x^2)$	$y' = \dfrac{7}{3}(2x)$	$y' = \dfrac{14x}{3}$
d.	$y = \dfrac{7}{(3x)^{-2}}$	$y = 63(x^2)$	$y' = 63(2x)$	$y' = 126x$

Try It **Exploration A**

The Sum and Difference Rules

THEOREM 2.5 The Sum and Difference Rules

The sum (or difference) of two differentiable functions f and g is itself differentiable. Moreover, the derivative of $f + g$ (or $f - g$) is the sum (or difference) of the derivatives of f and g.

$$\frac{d}{dx}[f(x) + g(x)] = f'(x) + g'(x) \qquad \text{Sum Rule}$$

$$\frac{d}{dx}[f(x) - g(x)] = f'(x) - g'(x) \qquad \text{Difference Rule}$$

Proof A proof of the Sum Rule follows from Theorem 1.2. (The Difference Rule can be proved in a similar way.)

$$\begin{aligned}
\frac{d}{dx}[f(x) + g(x)] &= \lim_{\Delta x \to 0} \frac{[f(x + \Delta x) + g(x + \Delta x)] - [f(x) + g(x)]}{\Delta x} \\
&= \lim_{\Delta x \to 0} \frac{f(x + \Delta x) + g(x + \Delta x) - f(x) - g(x)}{\Delta x} \\
&= \lim_{\Delta x \to 0} \left[\frac{f(x + \Delta x) - f(x)}{\Delta x} + \frac{g(x + \Delta x) - g(x)}{\Delta x}\right] \\
&= \lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x} + \lim_{\Delta x \to 0} \frac{g(x + \Delta x) - g(x)}{\Delta x} \\
&= f'(x) + g'(x)
\end{aligned}$$

The Sum and Difference Rules can be extended to any finite number of functions. For instance, if $F(x) = f(x) + g(x) - h(x)$, then $F'(x) = f'(x) + g'(x) - h'(x)$.

EXAMPLE 7 Using the Sum and Difference Rules

	Function	*Derivative*
a.	$f(x) = x^3 - 4x + 5$	$f'(x) = 3x^2 - 4$
b.	$g(x) = -\dfrac{x^4}{2} + 3x^3 - 2x$	$g'(x) = -2x^3 + 9x^2 - 2$

Try It **Exploration A**

The editable graph feature below allows you to edit the graph of a function and its derivative.

Editable Graph

FOR FURTHER INFORMATION For the outline of a geometric proof of the derivatives of the sine and cosine functions, see the article "The Spider's Spacewalk Derivation of sin′ and cos′" by Tim Hesterberg in *The College Mathematics Journal*.

MathArticle

Derivatives of Sine and Cosine Functions

In Section 1.3, you studied the following limits.

$$\lim_{\Delta x \to 0} \frac{\sin \Delta x}{\Delta x} = 1 \quad \text{and} \quad \lim_{\Delta x \to 0} \frac{1 - \cos \Delta x}{\Delta x} = 0$$

These two limits can be used to prove differentiation rules for the sine and cosine functions. (The derivatives of the other four trigonometric functions are discussed in Section 2.3.)

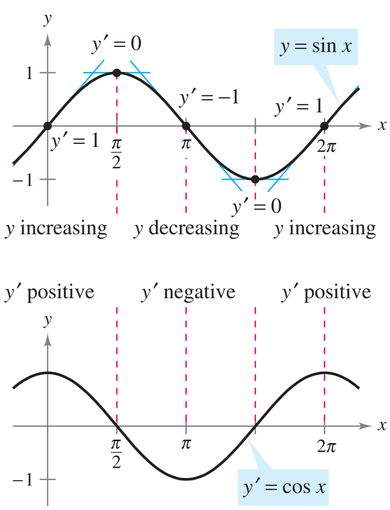
THEOREM 2.6 Derivatives of Sine and Cosine Functions

$$\frac{d}{dx}[\sin x] = \cos x \qquad \frac{d}{dx}[\cos x] = -\sin x$$

Proof

$$\begin{aligned}
\frac{d}{dx}[\sin x] &= \lim_{\Delta x \to 0} \frac{\sin(x + \Delta x) - \sin x}{\Delta x} && \text{Definition of derivative} \\
&= \lim_{\Delta x \to 0} \frac{\sin x \cos \Delta x + \cos x \sin \Delta x - \sin x}{\Delta x} \\
&= \lim_{\Delta x \to 0} \frac{\cos x \sin \Delta x - (\sin x)(1 - \cos \Delta x)}{\Delta x} \\
&= \lim_{\Delta x \to 0} \left[(\cos x)\left(\frac{\sin \Delta x}{\Delta x}\right) - (\sin x)\left(\frac{1 - \cos \Delta x}{\Delta x}\right) \right] \\
&= \cos x \left(\lim_{\Delta x \to 0} \frac{\sin \Delta x}{\Delta x} \right) - \sin x \left(\lim_{\Delta x \to 0} \frac{1 - \cos \Delta x}{\Delta x} \right) \\
&= (\cos x)(1) - (\sin x)(0) \\
&= \cos x
\end{aligned}$$

This differentiation rule is shown graphically in Figure 2.18. Note that for each x, the *slope* of the sine curve is equal to the value of the cosine. The proof of the second rule is left as an exercise (see Exercise 116).

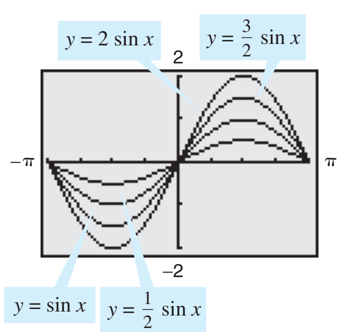


The derivative of the sine function is the cosine function.
Figure 2.18

Animation

EXAMPLE 8 Derivatives Involving Sines and Cosines

	Function	*Derivative*
a.	$y = 2 \sin x$	$y' = 2 \cos x$
b.	$y = \dfrac{\sin x}{2} = \dfrac{1}{2} \sin x$	$y' = \dfrac{1}{2} \cos x = \dfrac{\cos x}{2}$
c.	$y = x + \cos x$	$y' = 1 - \sin x$



$\frac{d}{dx}[a \sin x] = a \cos x$
Figure 2.19

TECHNOLOGY A graphing utility can provide insight into the interpretation of a derivative. For instance, Figure 2.19 shows the graphs of

$$y = a \sin x$$

for $a = \frac{1}{2}, 1, \frac{3}{2}$, and 2. Estimate the slope of each graph at the point $(0, 0)$. Then verify your estimates analytically by evaluating the derivative of each function when $x = 0$.

Try It | Exploration A | Open Exploration

Rates of Change

You have seen how the derivative is used to determine slope. The derivative can also be used to determine the rate of change of one variable with respect to another. Applications involving rates of change occur in a wide variety of fields. A few examples are population growth rates, production rates, water flow rates, velocity, and acceleration.

A common use for rate of change is to describe the motion of an object moving in a straight line. In such problems, it is customary to use either a horizontal or a vertical line with a designated origin to represent the line of motion. On such lines, movement to the right (or upward) is considered to be in the positive direction, and movement to the left (or downward) is considered to be in the negative direction.

The function s that gives the position (relative to the origin) of an object as a function of time t is called a **position function.** If, over a period of time Δt, the object changes its position by the amount $\Delta s = s(t + \Delta t) - s(t)$, then, by the familiar formula

$$\text{Rate} = \frac{\text{distance}}{\text{time}}$$

the **average velocity** is

$$\frac{\text{Change in distance}}{\text{Change in time}} = \frac{\Delta s}{\Delta t}.$$

Average velocity

EXAMPLE 9 Finding Average Velocity of a Falling Object

If a billiard ball is dropped from a height of 100 feet, its height s at time t is given by the position function

$$s = -16t^2 + 100$$

Position function

where s is measured in feet and t is measured in seconds. Find the average velocity over each of the following time intervals.

a. $[1, 2]$ **b.** $[1, 1.5]$ **c.** $[1, 1.1]$

Solution

a. For the interval $[1, 2]$, the object falls from a height of $s(1) = -16(1)^2 + 100 = 84$ feet to a height of $s(2) = -16(2)^2 + 100 = 36$ feet. The average velocity is

$$\frac{\Delta s}{\Delta t} = \frac{36 - 84}{2 - 1} = \frac{-48}{1} = -48 \text{ feet per second.}$$

b. For the interval $[1, 1.5]$, the object falls from a height of 84 feet to a height of 64 feet. The average velocity is

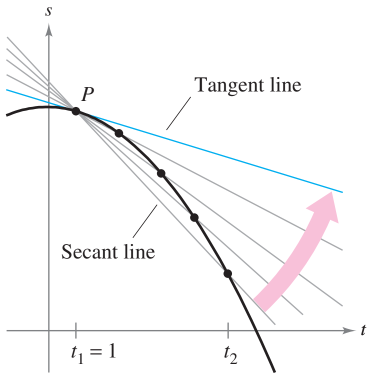
$$\frac{\Delta s}{\Delta t} = \frac{64 - 84}{1.5 - 1} = \frac{-20}{0.5} = -40 \text{ feet per second.}$$

c. For the interval $[1, 1.1]$, the object falls from a height of 84 feet to a height of 80.64 feet. The average velocity is

$$\frac{\Delta s}{\Delta t} = \frac{80.64 - 84}{1.1 - 1} = \frac{-3.36}{0.1} = -33.6 \text{ feet per second.}$$

Note that the average velocities are *negative*, indicating that the object is moving downward.

Try It Exploration A Exploration B



The average velocity between t_1 and t_2 is the slope of the secant line, and the instantaneous velocity at t_1 is the slope of the tangent line.
Figure 2.20

Animation

Suppose that in Example 9 you wanted to find the *instantaneous* velocity (or simply the velocity) of the object when $t = 1$. Just as you can approximate the slope of the tangent line by calculating the slope of the secant line, you can approximate the velocity at $t = 1$ by calculating the average velocity over a small interval $[1, 1 + \Delta t]$ (see Figure 2.20). By taking the limit as Δt approaches zero, you obtain the velocity when $t = 1$. Try doing this—you will find that the velocity when $t = 1$ is -32 feet per second.

In general, if $s = s(t)$ is the position function for an object moving along a straight line, the **velocity** of the object at time t is

$$v(t) = \lim_{\Delta t \to 0} \frac{s(t + \Delta t) - s(t)}{\Delta t} = s'(t).$$ Velocity function

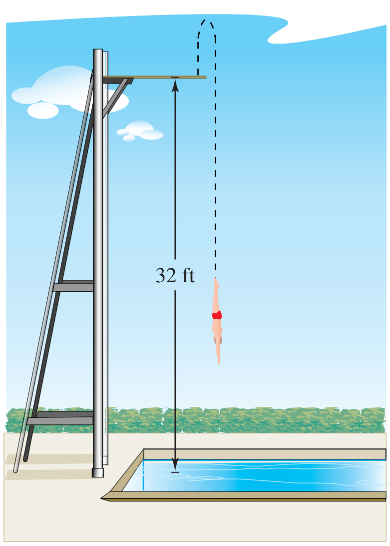
In other words, the velocity function is the derivative of the position function. Velocity can be negative, zero, or positive. The **speed** of an object is the absolute value of its velocity. Speed cannot be negative.

The position of a free-falling object (neglecting air resistance) under the influence of gravity can be represented by the equation

$$s(t) = \frac{1}{2}gt^2 + v_0 t + s_0$$ Position function

where s_0 is the initial height of the object, v_0 is the initial velocity of the object, and g is the acceleration due to gravity. On Earth, the value of g is approximately -32 feet per second per second or -9.8 meters per second per second.

History



Velocity is positive when an object is rising, and is negative when an object is falling.
Figure 2.21

Animation

NOTE In Figure 2.21, note that the diver moves upward for the first half-second because the velocity is positive for $0 < t < \frac{1}{2}$. When the velocity is 0, the diver has reached the maximum height of the dive.

EXAMPLE 10 Using the Derivative to Find Velocity

At time $t = 0$, a diver jumps from a platform diving board that is 32 feet above the water (see Figure 2.21). The position of the diver is given by

$$s(t) = -16t^2 + 16t + 32$$ Position function

where s is measured in feet and t is measured in seconds.

a. When does the diver hit the water?

b. What is the diver's velocity at impact?

Solution

a. To find the time t when the diver hits the water, let $s = 0$ and solve for t.

$$-16t^2 + 16t + 32 = 0$$ Set position function equal to 0.

$$-16(t + 1)(t - 2) = 0$$ Factor.

$$t = -1 \text{ or } 2$$ Solve for t.

Because $t \geq 0$, choose the positive value to conclude that the diver hits the water at $t = 2$ seconds.

b. The velocity at time t is given by the derivative $s'(t) = -32t + 16$. So, the velocity at time $t = 2$ is

$$s'(2) = -32(2) + 16 = -48 \text{ feet per second.}$$

Try It Exploration A

Section 2.3 Product and Quotient Rules and Higher-Order Derivatives

- **Find the derivative of a function using the Product Rule.**
- **Find the derivative of a function using the Quotient Rule.**
- **Find the derivative of a trigonometric function.**
- **Find a higher-order derivative of a function.**

The Product Rule

In Section 2.2 you learned that the derivative of the sum of two functions is simply the sum of their derivatives. The rules for the derivatives of the product and quotient of two functions are not as simple.

NOTE A version of the Product Rule that some people prefer is

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

The advantage of this form is that it generalizes easily to products involving three or more factors.

THEOREM 2.7 The Product Rule

The product of two differentiable functions f and g is itself differentiable. Moreover, the derivative of fg is the first function times the derivative of the second, plus the second function times the derivative of the first.

$$\frac{d}{dx}[f(x)g(x)] = f(x)g'(x) + g(x)f'(x)$$

Video

Proof Some mathematical proofs, such as the proof of the Sum Rule, are straightforward. Others involve clever steps that may appear unmotivated to a reader. This proof involves such a step—subtracting and adding the same quantity—which is shown in color.

$$\begin{aligned}
\frac{d}{dx}[f(x)g(x)] &= \lim_{\Delta x \to 0} \frac{f(x+\Delta x)g(x+\Delta x) - f(x)g(x)}{\Delta x} \\
&= \lim_{\Delta x \to 0} \frac{f(x+\Delta x)g(x+\Delta x) - f(x+\Delta x)g(x) + f(x+\Delta x)g(x) - f(x)g(x)}{\Delta x} \\
&= \lim_{\Delta x \to 0} \left[f(x+\Delta x)\frac{g(x+\Delta x) - g(x)}{\Delta x} + g(x)\frac{f(x+\Delta x) - f(x)}{\Delta x} \right] \\
&= \lim_{\Delta x \to 0} \left[f(x+\Delta x)\frac{g(x+\Delta x) - g(x)}{\Delta x} \right] + \lim_{\Delta x \to 0} \left[g(x)\frac{f(x+\Delta x) - f(x)}{\Delta x} \right] \\
&= \lim_{\Delta x \to 0} f(x+\Delta x) \cdot \lim_{\Delta x \to 0} \frac{g(x+\Delta x) - g(x)}{\Delta x} + \lim_{\Delta x \to 0} g(x) \cdot \lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x} \\
&= f(x)g'(x) + g(x)f'(x)
\end{aligned}$$

Note that $\lim\limits_{\Delta x \to 0} f(x + \Delta x) = f(x)$ because f is given to be differentiable and therefore is continuous.

The Product Rule can be extended to cover products involving more than two factors. For example, if f, g, and h are differentiable functions of x, then

$$\frac{d}{dx}[f(x)g(x)h(x)] = f'(x)g(x)h(x) + f(x)g'(x)h(x) + f(x)g(x)h'(x).$$

For instance, the derivative of $y = x^2 \sin x \cos x$ is

$$\begin{aligned}
\frac{dy}{dx} &= 2x \sin x \cos x + x^2 \cos x \cos x + x^2 \sin x(-\sin x) \\
&= 2x \sin x \cos x + x^2(\cos^2 x - \sin^2 x).
\end{aligned}$$

THE PRODUCT RULE

When Leibniz originally wrote a formula for the Product Rule, he was motivated by the expression

$$(x + dx)(y + dy) - xy$$

from which he subtracted $dx\,dy$ (as being negligible) and obtained the differential form $x\,dy + y\,dx$. This derivation resulted in the traditional form of the Product Rule. (*Source: The History of Mathematics by David M. Burton*)

The derivative of a product of two functions is not (in general) given by the product of the derivatives of the two functions. To see this, try comparing the product of the derivatives of $f(x) = 3x - 2x^2$ and $g(x) = 5 + 4x$ with the derivative in Example 1.

EXAMPLE 1 Using the Product Rule

Find the derivative of $h(x) = (3x - 2x^2)(5 + 4x)$.

Solution

$$\begin{aligned} h'(x) &= \overbrace{(3x - 2x^2)}^{\text{First}} \overbrace{\frac{d}{dx}[5 + 4x]}^{\text{Derivative of second}} + \overbrace{(5 + 4x)}^{\text{Second}} \overbrace{\frac{d}{dx}[3x - 2x^2]}^{\text{Derivative of first}} && \text{Apply Product Rule.} \\ &= (3x - 2x^2)(4) + (5 + 4x)(3 - 4x) \\ &= (12x - 8x^2) + (15 - 8x - 16x^2) \\ &= -24x^2 + 4x + 15 \end{aligned}$$

In Example 1, you have the option of finding the derivative with or without the Product Rule. To find the derivative without the Product Rule, you can write

$$\begin{aligned} D_x[(3x - 2x^2)(5 + 4x)] &= D_x[-8x^3 + 2x^2 + 15x] \\ &= -24x^2 + 4x + 15. \end{aligned}$$

Try It Exploration A

In the next example, you must use the Product Rule.

EXAMPLE 2 Using the Product Rule

Find the derivative of $y = 3x^2 \sin x$.

Solution

$$\begin{aligned} \frac{d}{dx}[3x^2 \sin x] &= 3x^2 \frac{d}{dx}[\sin x] + \sin x \frac{d}{dx}[3x^2] && \text{Apply Product Rule.} \\ &= 3x^2 \cos x + (\sin x)(6x) \\ &= 3x^2 \cos x + 6x \sin x \\ &= 3x(x \cos x + 2 \sin x) \end{aligned}$$

Try It Exploration A Technology

The editable graph feature below allows you to edit the graph of a function and its derivative.

Editable Graph

EXAMPLE 3 Using the Product Rule

Find the derivative of $y = 2x \cos x - 2 \sin x$.

Solution

$$\begin{aligned} \frac{dy}{dx} &= \overbrace{(2x)\left(\frac{d}{dx}[\cos x]\right) + (\cos x)\left(\frac{d}{dx}[2x]\right)}^{\text{Product Rule}} - \overbrace{2\frac{d}{dx}[\sin x]}^{\text{Constant Multiple Rule}} \\ &= (2x)(-\sin x) + (\cos x)(2) - 2(\cos x) \\ &= -2x \sin x \end{aligned}$$

NOTE In Example 3, notice that you use the Product Rule when both factors of the product are variable, and you use the Constant Multiple Rule when one of the factors is a constant.

Try It Exploration A Technology

The Quotient Rule

THEOREM 2.8 The Quotient Rule

The quotient f/g of two differentiable functions f and g is itself differentiable at all values of x for which $g(x) \neq 0$. Moreover, the derivative of f/g is given by the denominator times the derivative of the numerator minus the numerator times the derivative of the denominator, all divided by the square of the denominator.

$$\frac{d}{dx}\left[\frac{f(x)}{g(x)}\right] = \frac{g(x)f'(x) - f(x)g'(x)}{[g(x)]^2}, \qquad g(x) \neq 0$$

Video

Proof As with the proof of Theorem 2.7, the key to this proof is subtracting and adding the same quantity.

$$\frac{d}{dx}\left[\frac{f(x)}{g(x)}\right] = \lim_{\Delta x \to 0} \frac{\frac{f(x + \Delta x)}{g(x + \Delta x)} - \frac{f(x)}{g(x)}}{\Delta x} \qquad \text{Definition of derivative}$$

$$= \lim_{\Delta x \to 0} \frac{g(x)f(x + \Delta x) - f(x)g(x + \Delta x)}{\Delta x g(x)g(x + \Delta x)}$$

$$= \lim_{\Delta x \to 0} \frac{g(x)f(x + \Delta x) - f(x)g(x) + f(x)g(x) - f(x)g(x + \Delta x)}{\Delta x g(x)g(x + \Delta x)}$$

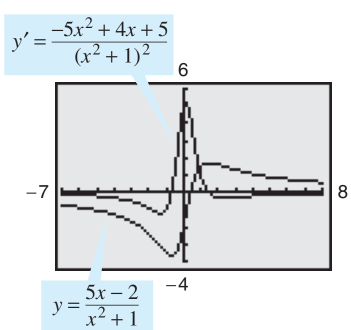
$$= \frac{\lim_{\Delta x \to 0} \frac{g(x)[f(x + \Delta x) - f(x)]}{\Delta x} - \lim_{\Delta x \to 0} \frac{f(x)[g(x + \Delta x) - g(x)]}{\Delta x}}{\lim_{\Delta x \to 0} [g(x)g(x + \Delta x)]}$$

$$= \frac{g(x)\left[\lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}\right] - f(x)\left[\lim_{\Delta x \to 0} \frac{g(x + \Delta x) - g(x)}{\Delta x}\right]}{\lim_{\Delta x \to 0} [g(x)g(x + \Delta x)]}$$

$$= \frac{g(x)f'(x) - f(x)g'(x)}{[g(x)]^2}$$

Note that $\lim_{\Delta x \to 0} g(x + \Delta x) = g(x)$ because g is given to be differentiable and therefore is continuous.

TECHNOLOGY A graphing utility can be used to compare the graph of a function with the graph of its derivative. For instance, in Figure 2.22, the graph of the function in Example 4 appears to have two points that have horizontal tangent lines. What are the values of y' at these two points?



Graphical comparison of a function and its derivative
Figure 2.22

EXAMPLE 4 Using the Quotient Rule

Find the derivative of $y = \frac{5x - 2}{x^2 + 1}$.

Solution

$$\frac{d}{dx}\left[\frac{5x - 2}{x^2 + 1}\right] = \frac{(x^2 + 1)\frac{d}{dx}[5x - 2] - (5x - 2)\frac{d}{dx}[x^2 + 1]}{(x^2 + 1)^2} \qquad \text{Apply Quotient Rule.}$$

$$= \frac{(x^2 + 1)(5) - (5x - 2)(2x)}{(x^2 + 1)^2}$$

$$= \frac{(5x^2 + 5) - (10x^2 - 4x)}{(x^2 + 1)^2}$$

$$= \frac{-5x^2 + 4x + 5}{(x^2 + 1)^2}$$

Try It Exploration A Exploration B

The editable graph feature below allows you to edit the graph of a function and its derivative.

Editable Graph

Note the use of parentheses in Example 4. A liberal use of parentheses is recommended for *all* types of differentiation problems. For instance, with the Quotient Rule, it is a good idea to enclose all factors and derivatives in parentheses, and to pay special attention to the subtraction required in the numerator.

When differentiation rules were introduced in the preceding section, the need for rewriting *before* differentiating was emphasized. The next example illustrates this point with the Quotient Rule.

EXAMPLE 5 Rewriting Before Differentiating

Find an equation of the tangent line to the graph of $f(x) = \dfrac{3 - (1/x)}{x + 5}$ at $(-1, 1)$.

Solution Begin by rewriting the function.

$$f(x) = \frac{3 - (1/x)}{x + 5} \qquad \text{Write original function.}$$

$$= \frac{x\left(3 - \frac{1}{x}\right)}{x(x + 5)} \qquad \text{Multiply numerator and denominator by } x.$$

$$= \frac{3x - 1}{x^2 + 5x} \qquad \text{Rewrite.}$$

$$f'(x) = \frac{(x^2 + 5x)(3) - (3x - 1)(2x + 5)}{(x^2 + 5x)^2} \qquad \text{Quotient Rule}$$

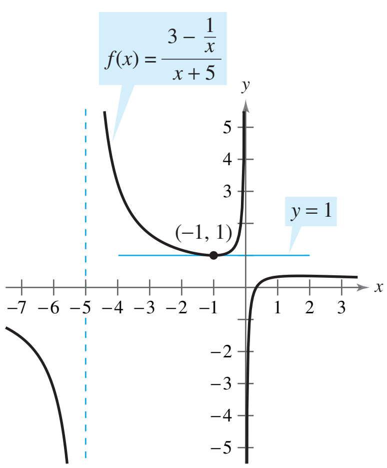
$$= \frac{(3x^2 + 15x) - (6x^2 + 13x - 5)}{(x^2 + 5x)^2}$$

$$= \frac{-3x^2 + 2x + 5}{(x^2 + 5x)^2} \qquad \text{Simplify.}$$

To find the slope at $(-1, 1)$, evaluate $f'(-1)$.

$$f'(-1) = 0 \qquad \text{Slope of graph at } (-1, 1)$$

Then, using the point-slope form of the equation of a line, you can determine that the equation of the tangent line at $(-1, 1)$ is $y = 1$. See Figure 2.23.



The line $y = 1$ is tangent to the graph of $f(x)$ at the point $(-1, 1)$.
Figure 2.23

Try It | Exploration A

The editable graph feature below allows you to edit the graph of a function.

Editable Graph

Not every quotient needs to be differentiated by the Quotient Rule. For example, each quotient in the next example can be considered as the product of a constant times a function of x. In such cases it is more convenient to use the Constant Multiple Rule.

EXAMPLE 6 Using the Constant Multiple Rule

	Original Function	*Rewrite*	*Differentiate*	*Simplify*
a.	$y = \dfrac{x^2 + 3x}{6}$	$y = \dfrac{1}{6}(x^2 + 3x)$	$y' = \dfrac{1}{6}(2x + 3)$	$y' = \dfrac{2x + 3}{6}$
b.	$y = \dfrac{5x^4}{8}$	$y = \dfrac{5}{8}x^4$	$y' = \dfrac{5}{8}(4x^3)$	$y' = \dfrac{5}{2}x^3$
c.	$y = \dfrac{-3(3x - 2x^2)}{7x}$	$y = -\dfrac{3}{7}(3 - 2x)$	$y' = -\dfrac{3}{7}(-2)$	$y' = \dfrac{6}{7}$
d.	$y = \dfrac{9}{5x^2}$	$y = \dfrac{9}{5}(x^{-2})$	$y' = \dfrac{9}{5}(-2x^{-3})$	$y' = -\dfrac{18}{5x^3}$

NOTE To see the benefit of using the Constant Multiple Rule for some quotients, try using the Quotient Rule to differentiate the functions in Example 6—you should obtain the same results, but with more work.

Try It | Exploration A

In Section 2.2, the Power Rule was proved only for the case where the exponent n is a positive integer greater than 1. The next example extends the proof to include negative integer exponents.

EXAMPLE 7 Proof of the Power Rule (Negative Integer Exponents)

If n is a negative integer, there exists a positive integer k such that $n = -k$. So, by the Quotient Rule, you can write

$$\begin{aligned}
\frac{d}{dx}[x^n] &= \frac{d}{dx}\left[\frac{1}{x^k}\right] \\
&= \frac{x^k(0) - (1)(kx^{k-1})}{(x^k)^2} && \text{Quotient Rule and Power Rule} \\
&= \frac{0 - kx^{k-1}}{x^{2k}} \\
&= -kx^{-k-1} \\
&= nx^{n-1}. && n = -k
\end{aligned}$$

So, the Power Rule

$$D_x[x^n] = nx^{n-1} \qquad \text{Power Rule}$$

is valid for any integer. In Exercise 75 in Section 2.5, you are asked to prove the case for which n is any rational number.

Try It

Exploration A

Derivatives of Trigonometric Functions

Knowing the derivatives of the sine and cosine functions, you can use the Quotient Rule to find the derivatives of the four remaining trigonometric functions.

THEOREM 2.9 Derivatives of Trigonometric Functions

$$\frac{d}{dx}[\tan x] = \sec^2 x \qquad \frac{d}{dx}[\cot x] = -\csc^2 x$$

$$\frac{d}{dx}[\sec x] = \sec x \tan x \qquad \frac{d}{dx}[\csc x] = -\csc x \cot x$$

Video

Proof Considering $\tan x = (\sin x)/(\cos x)$ and applying the Quotient Rule, you obtain

$$\begin{aligned}
\frac{d}{dx}[\tan x] &= \frac{(\cos x)(\cos x) - (\sin x)(-\sin x)}{\cos^2 x} && \text{Apply Quotient Rule.} \\
&= \frac{\cos^2 x + \sin^2 x}{\cos^2 x} \\
&= \frac{1}{\cos^2 x} \\
&= \sec^2 x.
\end{aligned}$$

The proofs of the other three parts of the theorem are left as an exercise (see Exercise 89).

NOTE Because of trigonometric identities, the derivative of a trigonometric function can take many forms. This presents a challenge when you are trying to match your answers to those given in the back of the text.

EXAMPLE 8 Differentiating Trigonometric Functions

Function	*Derivative*
a. $y = x - \tan x$	$\dfrac{dy}{dx} = 1 - \sec^2 x$
b. $y = x \sec x$	$y' = x(\sec x \tan x) + (\sec x)(1)$
	$= (\sec x)(1 + x \tan x)$

Try It | Exploration A | Open Exploration

EXAMPLE 9 Different Forms of a Derivative

Differentiate both forms of $y = \dfrac{1 - \cos x}{\sin x} = \csc x - \cot x.$

Solution

First form: $y = \dfrac{1 - \cos x}{\sin x}$

$$y' = \frac{(\sin x)(\sin x) - (1 - \cos x)(\cos x)}{\sin^2 x}$$

$$= \frac{\sin^2 x + \cos^2 x - \cos x}{\sin^2 x}$$

$$= \frac{1 - \cos x}{\sin^2 x}$$

Second form: $y = \csc x - \cot x$

$$y' = -\csc x \cot x + \csc^2 x$$

To show that the two derivatives are equal, you can write

$$\frac{1 - \cos x}{\sin^2 x} = \frac{1}{\sin^2 x} - \left(\frac{1}{\sin x}\right)\left(\frac{\cos x}{\sin x}\right)$$

$$= \csc^2 x - \csc x \cot x.$$

Try It | Exploration A | Technology

The summary below shows that much of the work in obtaining a simplified form of a derivative occurs *after* differentiating. Note that two characteristics of a simplified form are the absence of negative exponents and the combining of like terms.

	$f'(x)$ After Differentiating	*$f'(x)$ After Simplifying*
Example 1	$(3x - 2x^2)(4) + (5 + 4x)(3 - 4x)$	$-24x^2 + 4x + 15$
Example 3	$(2x)(-\sin x) + (\cos x)(2) - 2(\cos x)$	$-2x \sin x$
Example 4	$\dfrac{(x^2 + 1)(5) - (5x - 2)(2x)}{(x^2 + 1)^2}$	$\dfrac{-5x^2 + 4x + 5}{(x^2 + 1)^2}$
Example 5	$\dfrac{(x^2 + 5x)(3) - (3x - 1)(2x + 5)}{(x^2 + 5x)^2}$	$\dfrac{-3x^2 + 2x + 5}{(x^2 + 5x)^2}$
Example 9	$\dfrac{(\sin x)(\sin x) - (1 - \cos x)(\cos x)}{\sin^2 x}$	$\dfrac{1 - \cos x}{\sin^2 x}$

Higher-Order Derivatives

Just as you can obtain a velocity function by differentiating a position function, you can obtain an **acceleration** function by differentiating a velocity function. Another way of looking at this is that you can obtain an acceleration function by differentiating a position function *twice*.

$$s(t) \qquad \text{Position function}$$
$$v(t) = s'(t) \qquad \text{Velocity function}$$
$$a(t) = v'(t) = s''(t) \qquad \text{Acceleration function}$$

NOTE: The second derivative of f is the derivative of the first derivative of f.

The function given by $a(t)$ is the **second derivative** of $s(t)$ and is denoted by $s''(t)$.

The second derivative is an example of a **higher-order derivative.** You can define derivatives of any positive integer order. For instance, the **third derivative** is the derivative of the second derivative. Higher-order derivatives are denoted as follows.

First derivative:	y',	$f'(x)$,	$\dfrac{dy}{dx}$,	$\dfrac{d}{dx}[f(x)]$,	$D_x[y]$
Second derivative:	y'',	$f''(x)$,	$\dfrac{d^2y}{dx^2}$,	$\dfrac{d^2}{dx^2}[f(x)]$,	$D_x^2[y]$
Third derivative:	y''',	$f'''(x)$,	$\dfrac{d^3y}{dx^3}$,	$\dfrac{d^3}{dx^3}[f(x)]$,	$D_x^3[y]$
Fourth derivative:	$y^{(4)}$,	$f^{(4)}(x)$,	$\dfrac{d^4y}{dx^4}$,	$\dfrac{d^4}{dx^4}[f(x)]$,	$D_x^4[y]$
$\vdots$					
nth derivative:	$y^{(n)}$,	$f^{(n)}(x)$,	$\dfrac{d^ny}{dx^n}$,	$\dfrac{d^n}{dx^n}[f(x)]$,	$D_x^n[y]$

EXAMPLE 10 Finding the Acceleration Due to Gravity

Because the moon has no atmosphere, a falling object on the moon encounters no air resistance. In 1971, astronaut David Scott demonstrated that a feather and a hammer fall at the same rate on the moon. The position function for each of these falling objects is given by

$$s(t) = -0.81t^2 + 2$$

where $s(t)$ is the height in meters and t is the time in seconds. What is the ratio of Earth's gravitational force to the moon's?

Solution To find the acceleration, differentiate the position function twice.

$$s(t) = -0.81t^2 + 2 \qquad \text{Position function}$$
$$s'(t) = -1.62t \qquad \text{Velocity function}$$
$$s''(t) = -1.62 \qquad \text{Acceleration function}$$

So, the acceleration due to gravity on the moon is -1.62 meters per second per second. Because the acceleration due to gravity on Earth is -9.8 meters per second per second, the ratio of Earth's gravitational force to the moon's is

$$\frac{\text{Earth's gravitational force}}{\text{Moon's gravitational force}} = \frac{-9.8}{-1.62} \approx 6.05.$$

THE MOON

The moon's mass is 7.349×10^{22} kilograms, and Earth's mass is 5.976×10^{24} kilograms. The moon's radius is 1737 kilometers, and Earth's radius is 6378 kilometers. Because the gravitational force on the surface of a planet is directly proportional to its mass and inversely proportional to the square of its radius, the ratio of the gravitational force on Earth to the gravitational force on the moon is

$$\frac{(5.976 \times 10^{24})/6378^2}{(7.349 \times 10^{22})/1737^2} \approx 6.03.$$

Video

Try It

Exploration A

Section 2.4 The Chain Rule

- **Find the derivative of a composite function using the Chain Rule.**
- **Find the derivative of a function using the General Power Rule.**
- **Simplify the derivative of a function using algebra.**
- **Find the derivative of a trigonometric function using the Chain Rule.**

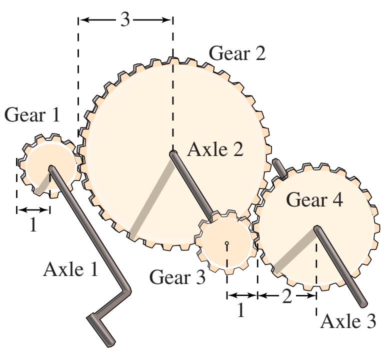
The Chain Rule

This text has yet to discuss one of the most powerful differentiation rules—the **Chain Rule.** This rule deals with composite functions and adds a surprising versatility to the rules discussed in the two previous sections. For example, compare the functions shown below. Those on the left can be differentiated without the Chain Rule, and those on the right are best done with the Chain Rule.

Without the Chain Rule	*With the Chain Rule*
$y = x^2 + 1$	$y = \sqrt{x^2 + 1}$
$y = \sin x$	$y = \sin 6x$
$y = 3x + 2$	$y = (3x + 2)^5$
$y = x + \tan x$	$y = x + \tan x^2$

Basically, the Chain Rule states that if y changes dy/du times as fast as u, and u changes du/dx times as fast as x, then y changes $(dy/du)(du/dx)$ times as fast as x.

Video



Axle 1: y revolutions per minute
Axle 2: u revolutions per minute
Axle 3: x revolutions per minute
Figure 2.24

Animation

EXAMPLE 1 The Derivative of a Composite Function

A set of gears is constructed, as shown in Figure 2.24, such that the second and third gears are on the same axle. As the first axle revolves, it drives the second axle, which in turn drives the third axle. Let y, u, and x represent the numbers of revolutions per minute of the first, second, and third axles. Find dy/du, du/dx, and dy/dx, and show that

$$\frac{dy}{dx} = \frac{dy}{du} \cdot \frac{du}{dx}.$$

Solution Because the circumference of the second gear is three times that of the first, the first axle must make three revolutions to turn the second axle once. Similarly, the second axle must make two revolutions to turn the third axle once, and you can write

$$\frac{dy}{du} = 3 \quad \text{and} \quad \frac{du}{dx} = 2.$$

Combining these two results, you know that the first axle must make six revolutions to turn the third axle once. So, you can write

$$\begin{aligned} \frac{dy}{dx} &= \boxed{\begin{array}{l}\text{Rate of change of first axle}\\ \text{with respect to second axle}\end{array}} \cdot \boxed{\begin{array}{l}\text{Rate of change of second axle}\\ \text{with respect to third axle}\end{array}} \\ &= \frac{dy}{du} \cdot \frac{du}{dx} = 3 \cdot 2 = 6 \\ &= \boxed{\begin{array}{l}\text{Rate of change of first axle}\\ \text{with respect to third axle}\end{array}}. \end{aligned}$$

In other words, the rate of change of y with respect to x is the product of the rate of change of y with respect to u and the rate of change of u with respect to x.

Try It Exploration A

EXPLORATION

Using the Chain Rule Each of the following functions can be differentiated using rules that you studied in Sections 2.2 and 2.3. For each function, find the derivative using those rules. Then find the derivative using the Chain Rule. Compare your results. Which method is simpler?

a. $\dfrac{2}{3x + 1}$

b. $(x + 2)^3$

c. $\sin 2x$

Example 1 illustrates a simple case of the Chain Rule. The general rule is stated below.

THEOREM 2.10 The Chain Rule

If $y = f(u)$ is a differentiable function of u and $u = g(x)$ is a differentiable function of x, then $y = f(g(x))$ is a differentiable function of x and

$$\frac{dy}{dx} = \frac{dy}{du} \cdot \frac{du}{dx}$$

or, equivalently,

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

Proof Let $h(x) = f(g(x))$. Then, using the alternative form of the derivative, you need to show that, for $x = c$,

$$h'(c) = f'(g(c))g'(c).$$

An important consideration in this proof is the behavior of g as x approaches c. A problem occurs if there are values of x, other than c, such that $g(x) = g(c)$. Appendix A shows how to use the differentiability of f and g to overcome this problem. For now, assume that $g(x) \neq g(c)$ for values of x other than c. In the proofs of the Product Rule and the Quotient Rule, the same quantity was added and subtracted to obtain the desired form. This proof uses a similar technique—multiplying and dividing by the same (nonzero) quantity. Note that because g is differentiable, it is also continuous, and it follows that $g(x) \to g(c)$ as $x \to c$.

$$\begin{aligned}
h'(c) &= \lim_{x\to c} \frac{f(g(x)) - f(g(c))}{x - c} \\
&= \lim_{x\to c} \left[\frac{f(g(x)) - f(g(c))}{g(x) - g(c)} \cdot \frac{g(x) - g(c)}{x - c}\right], \quad g(x) \neq g(c) \\
&= \left[\lim_{x\to c} \frac{f(g(x)) - f(g(c))}{g(x) - g(c)}\right]\left[\lim_{x\to c} \frac{g(x) - g(c)}{x - c}\right] \\
&= f'(g(c))g'(c)
\end{aligned}$$

When applying the Chain Rule, it is helpful to think of the composite function $f \circ g$ as having two parts—an inner part and an outer part.

Outer function

$$y = f(g(x)) = f(u)$$

Inner function

The derivative of $y = f(u)$ is the derivative of the outer function (at the inner function u) *times* the derivative of the inner function.

$$y' = f'(u) \cdot u'$$

EXAMPLE 2 Decomposition of a Composite Function

$y = f(g(x))$	$u = g(x)$	$y = f(u)$
a. $y = \dfrac{1}{x+1}$	$u = x + 1$	$y = \dfrac{1}{u}$
b. $y = \sin 2x$	$u = 2x$	$y = \sin u$
c. $y = \sqrt{3x^2 - x + 1}$	$u = 3x^2 - x + 1$	$y = \sqrt{u}$
d. $y = \tan^2 x$	$u = \tan x$	$y = u^2$

Try It

Exploration A

EXAMPLE 3 Using the Chain Rule

STUDY TIP You could also solve the problem in Example 3 without using the Chain Rule by observing that

$$y = x^6 + 3x^4 + 3x^2 + 1$$

and

$$y' = 6x^5 + 12x^3 + 6x.$$

Verify that this is the same as the derivative in Example 3. Which method would you use to find

$$\frac{d}{dx}(x^2 + 1)^{50}?$$

Find dy/dx for $y = (x^2 + 1)^3$.

Solution For this function, you can consider the inside function to be $u = x^2 + 1$. By the Chain Rule, you obtain

$$\frac{dy}{dx} = \underbrace{3(x^2 + 1)^2}_{\frac{dy}{du}}\underbrace{(2x)}_{\frac{du}{dx}} = 6x(x^2 + 1)^2.$$

Try It

Exploration A

Exploration B

The editable graph feature below allows you to edit the graph of a function and its derivative.

Editable Graph

The General Power Rule

The function in Example 3 is an example of one of the most common types of composite functions, $y = [u(x)]^n$. The rule for differentiating such functions is called the **General Power Rule,** and it is a special case of the Chain Rule.

THEOREM 2.11 The General Power Rule

If $y = [u(x)]^n$, where u is a differentiable function of x and n is a rational number, then

$$\frac{dy}{dx} = n[u(x)]^{n-1}\frac{du}{dx}$$

or, equivalently,

$$\frac{d}{dx}[u^n] = n u^{n-1} u'.$$

Video

Proof Because $y = u^n$, you apply the Chain Rule to obtain

$$\frac{dy}{dx} = \left(\frac{dy}{du}\right)\left(\frac{du}{dx}\right)$$
$$= \frac{d}{du}[u^n]\frac{du}{dx}.$$

By the (Simple) Power Rule in Section 2.2, you have $D_u[u^n] = nu^{n-1}$, and it follows that

$$\frac{dy}{dx} = n[u(x)]^{n-1}\frac{du}{dx}.$$

Video

EXAMPLE 4 Applying the General Power Rule

Find the derivative of $f(x) = (3x - 2x^2)^3$.

Solution Let $u = 3x - 2x^2$. Then

$$f(x) = (3x - 2x^2)^3 = u^3$$

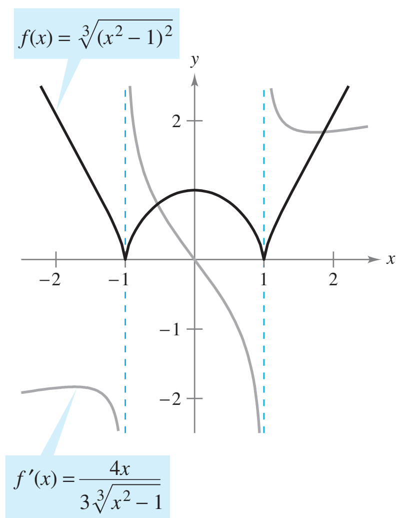
and, by the General Power Rule, the derivative is

$$\begin{aligned} f'(x) &= \overset{n}{3}\overbrace{(3x - 2x^2)^2}^{u^{n-1}}\overbrace{\frac{d}{dx}[3x - 2x^2]}^{u'} && \text{Apply General Power Rule.} \\ &= 3(3x - 2x^2)^2(3 - 4x). && \text{Differentiate } 3x - 2x^2. \end{aligned}$$

Try It Exploration A

The editable graph feature below allows you to edit the graph of a function.

Editable Graph



The derivative of f is 0 at $x = 0$ and is undefined at $x = \pm 1$.
Figure 2.25

Editable Graph

EXAMPLE 5 Differentiating Functions Involving Radicals

Find all points on the graph of $f(x) = \sqrt[3]{(x^2 - 1)^2}$ for which $f'(x) = 0$ and those for which $f'(x)$ does not exist.

Solution Begin by rewriting the function as

$$f(x) = (x^2 - 1)^{2/3}.$$

Then, applying the General Power Rule (with $u = x^2 - 1$) produces

$$\begin{aligned} f'(x) &= \overset{n}{\frac{2}{3}}\overbrace{(x^2 - 1)^{-1/3}}^{u^{n-1}}\overbrace{(2x)}^{u'} && \text{Apply General Power Rule.} \\ &= \frac{4x}{3\sqrt[3]{x^2 - 1}}. && \text{Write in radical form.} \end{aligned}$$

So, $f'(x) = 0$ when $x = 0$ and $f'(x)$ does not exist when $x = \pm 1$, as shown in Figure 2.25.

Try It Exploration A

EXAMPLE 6 Differentiating Quotients with Constant Numerators

Differentiate $g(t) = \dfrac{-7}{(2t - 3)^2}$.

Solution Begin by rewriting the function as

$$g(t) = -7(2t - 3)^{-2}.$$

Then, applying the General Power Rule produces

$$\begin{aligned} g'(t) &= \underbrace{(-7)}_{\substack{\text{Constant} \\ \text{Multiple Rule}}}\overbrace{(-2)}^{n}\overbrace{(2t - 3)^{-3}}^{u^{n-1}}\overbrace{(2)}^{u'} && \text{Apply General Power Rule.} \\ &= 28(2t - 3)^{-3} && \text{Simplify.} \\ &= \frac{28}{(2t - 3)^3}. && \text{Write with positive exponent.} \end{aligned}$$

NOTE Try differentiating the function in Example 6 using the Quotient Rule. You should obtain the same result, but using the Quotient Rule is less efficient than using the General Power Rule.

Try It Exploration A Exploration B

Simplifying Derivatives

The next three examples illustrate some techniques for simplifying the "raw derivatives" of functions involving products, quotients, and composites.

EXAMPLE 7 Simplifying by Factoring Out the Least Powers

$$
\begin{aligned}
f(x) &= x^2\sqrt{1-x^2} && \text{Original function}\\
&= x^2(1-x^2)^{1/2} && \text{Rewrite.}\\
f'(x) &= x^2\frac{d}{dx}\left[(1-x^2)^{1/2}\right] + (1-x^2)^{1/2}\frac{d}{dx}\left[x^2\right] && \text{Product Rule}\\
&= x^2\left[\frac{1}{2}(1-x^2)^{-1/2}(-2x)\right] + (1-x^2)^{1/2}(2x) && \text{General Power Rule}\\
&= -x^3(1-x^2)^{-1/2} + 2x(1-x^2)^{1/2} && \text{Simplify.}\\
&= x(1-x^2)^{-1/2}\left[-x^2(1) + 2(1-x^2)\right] && \text{Factor.}\\
&= \frac{x(2-3x^2)}{\sqrt{1-x^2}} && \text{Simplify.}
\end{aligned}
$$

Try It Exploration A

EXAMPLE 8 Simplifying the Derivative of a Quotient

TECHNOLOGY Symbolic differentiation utilities are capable of differentiating very complicated functions. Often, however, the result is given in unsimplified form. If you have access to such a utility, use it to find the derivatives of the functions given in Examples 7, 8, and 9. Then compare the results with those given on this page.

$$
\begin{aligned}
f(x) &= \frac{x}{\sqrt[3]{x^2+4}} && \text{Original function}\\
&= \frac{x}{(x^2+4)^{1/3}} && \text{Rewrite.}\\
f'(x) &= \frac{(x^2+4)^{1/3}(1) - x(1/3)(x^2+4)^{-2/3}(2x)}{(x^2+4)^{2/3}} && \text{Quotient Rule}\\
&= \frac{1}{3}(x^2+4)^{-2/3}\left[\frac{3(x^2+4) - (2x^2)(1)}{(x^2+4)^{2/3}}\right] && \text{Factor.}\\
&= \frac{x^2+12}{3(x^2+4)^{4/3}} && \text{Simplify.}
\end{aligned}
$$

Try It Exploration A

EXAMPLE 9 Simplifying the Derivative of a Power

$$
\begin{aligned}
y &= \left(\frac{3x-1}{x^2+3}\right)^2 && \text{Original function}\\
y' &= \overset{n}{2}\overbrace{\left(\frac{3x-1}{x^2+3}\right)}^{u^{n-1}}\overbrace{\frac{d}{dx}\left[\frac{3x-1}{x^2+3}\right]}^{u'} && \text{General Power Rule}\\
&= \left[\frac{2(3x-1)}{x^2+3}\right]\left[\frac{(x^2+3)(3) - (3x-1)(2x)}{(x^2+3)^2}\right] && \text{Quotient Rule}\\
&= \frac{2(3x-1)(3x^2+9-6x^2+2x)}{(x^2+3)^3} && \text{Multiply.}\\
&= \frac{2(3x-1)(-3x^2+2x+9)}{(x^2+3)^3} && \text{Simplify.}
\end{aligned}
$$

Try It Exploration A Open Exploration

Trigonometric Functions and the Chain Rule

The "Chain Rule versions" of the derivatives of the six trigonometric functions are as shown.

$$\frac{d}{dx}[\sin u] = (\cos u)\,u' \qquad \frac{d}{dx}[\cos u] = -(\sin u)\,u'$$

$$\frac{d}{dx}[\tan u] = (\sec^2 u)\,u' \qquad \frac{d}{dx}[\cot u] = -(\csc^2 u)\,u'$$

$$\frac{d}{dx}[\sec u] = (\sec u \tan u)\,u' \qquad \frac{d}{dx}[\csc u] = -(\csc u \cot u)\,u'$$

Technology

EXAMPLE 10 Applying the Chain Rule to Trigonometric Functions

a. $y = \sin \overbrace{2x}^{u}$ $\qquad y' = \overbrace{\cos 2x}^{\cos u}\,\overbrace{\frac{d}{dx}[2x]}^{u'} = (\cos 2x)(2) = 2\cos 2x$

b. $y = \cos(x - 1)$ $\qquad y' = -\sin(x - 1)$

c. $y = \tan 3x$ $\qquad y' = 3\sec^2 3x$

Try It Exploration A

Be sure that you understand the mathematical conventions regarding parentheses and trigonometric functions. For instance, in Example 10(a), $\sin 2x$ is written to mean $\sin(2x)$.

EXAMPLE 11 Parentheses and Trigonometric Functions

a. $y = \cos 3x^2 = \cos(3x^2)$ $\qquad y' = (-\sin 3x^2)(6x) = -6x \sin 3x^2$

b. $y = (\cos 3)x^2$ $\qquad y' = (\cos 3)(2x) = 2x \cos 3$

c. $y = \cos(3x)^2 = \cos(9x^2)$ $\qquad y' = (-\sin 9x^2)(18x) = -18x \sin 9x^2$

d. $y = \cos^2 x = (\cos x)^2$ $\qquad y' = 2(\cos x)(-\sin x) = -2\cos x \sin x$

e. $y = \sqrt{\cos x} = (\cos x)^{1/2}$ $\qquad y' = \frac{1}{2}(\cos x)^{-1/2}(-\sin x) = -\frac{\sin x}{2\sqrt{\cos x}}$

Try It Exploration A

To find the derivative of a function of the form $k(x) = f(g(h(x)))$, you need to apply the Chain Rule twice, as shown in Example 12.

EXAMPLE 12 Repeated Application of the Chain Rule

$$\begin{aligned}
f(t) &= \sin^3 4t && \text{Original function} \\
&= (\sin 4t)^3 && \text{Rewrite.} \\
f'(t) &= 3(\sin 4t)^2 \frac{d}{dt}[\sin 4t] && \text{Apply Chain Rule once.} \\
&= 3(\sin 4t)^2(\cos 4t)\frac{d}{dt}[4t] && \text{Apply Chain Rule a second time.} \\
&= 3(\sin 4t)^2(\cos 4t)(4) && \\
&= 12 \sin^2 4t \cos 4t && \text{Simplify.}
\end{aligned}$$

Try It Exploration A

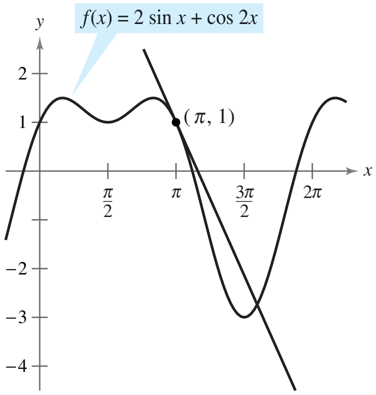


Figure 2.26

EXAMPLE 13 Tangent Line of a Trigonometric Function

Find an equation of the tangent line to the graph of

$$f(x) = 2 \sin x + \cos 2x$$

at the point $(\pi, 1)$, as shown in Figure 2.26. Then determine all values of x in the interval $(0, 2\pi)$ at which the graph of f has a horizontal tangent.

Solution Begin by finding $f'(x)$.

$$f(x) = 2 \sin x + \cos 2x \qquad \text{Write original function.}$$

$$f'(x) = 2 \cos x + (-\sin 2x)(2) \qquad \text{Apply Chain Rule to } \cos 2x.$$

$$= 2 \cos x - 2 \sin 2x \qquad \text{Simplify.}$$

To find the equation of the tangent line at $(\pi, 1)$, evaluate $f'(\pi)$.

$$f'(\pi) = 2 \cos \pi - 2 \sin 2\pi \qquad \text{Substitute.}$$

$$= -2 \qquad \text{Slope of graph at } (\pi, 1)$$

Now, using the point-slope form of the equation of a line, you can write

$$y - y_1 = m(x - x_1) \qquad \text{Point-slope form}$$

$$y - 1 = -2(x - \pi) \qquad \text{Substitute for } y_1, m, \text{ and } x_1.$$

$$y = 1 - 2x + 2\pi. \qquad \text{Equation of tangent line at } (\pi, 1)$$

You can then determine that $f'(x) = 0$ when $x = \frac{\pi}{6}, \frac{\pi}{2}, \frac{5\pi}{6}$, and $\frac{3\pi}{2}$. So, f has a horizontal tangent at $x = \frac{\pi}{6}, \frac{\pi}{2}, \frac{5\pi}{6}$, and $\frac{3\pi}{2}$.

STUDY TIP To become skilled at differentiation, you should memorize each rule. As an aid to memorization, note that the cofunctions (cosine, cotangent, and cosecant) require a negative sign as part of their derivatives.

Try It   Exploration A

This section concludes with a summary of the differentiation rules studied so far.

Summary of Differentiation Rules

General Differentiation Rules Let f, g, and u be differentiable functions of x.

Constant Multiple Rule: $\frac{d}{dx}[cf] = cf'$

Sum or Difference Rule: $\frac{d}{dx}[f \pm g] = f' \pm g'$

Product Rule: $\frac{d}{dx}[fg] = fg' + gf'$

Quotient Rule: $\frac{d}{dx}\left[\frac{f}{g}\right] = \frac{gf' - fg'}{g^2}$

Derivatives of Algebraic Functions

Constant Rule: $\frac{d}{dx}[c] = 0$

(Simple) Power Rule: $\frac{d}{dx}[x^n] = nx^{n-1}$, $\frac{d}{dx}[x] = 1$

Derivatives of Trigonometric Functions

$\frac{d}{dx}[\sin x] = \cos x$ $\quad \frac{d}{dx}[\tan x] = \sec^2 x$ $\quad \frac{d}{dx}[\sec x] = \sec x \tan x$

$\frac{d}{dx}[\cos x] = -\sin x$ $\quad \frac{d}{dx}[\cot x] = -\csc^2 x$ $\quad \frac{d}{dx}[\csc x] = -\csc x \cot x$

Chain Rule

Chain Rule: $\frac{d}{dx}[f(u)] = f'(u)\, u'$

General Power Rule: $\frac{d}{dx}[u^n] = nu^{n-1}\, u'$

Section 2.5 Implicit Differentiation

- **Distinguish between functions written in implicit form and explicit form.**
- **Use implicit differentiation to find the derivative of a function.**

EXPLORATION

Graphing an Implicit Equation
How could you use a graphing utility to sketch the graph of the equation

$$x^2 - 2y^3 + 4y = 2?$$

Here are two possible approaches.

a. Solve the equation for x. Switch the roles of x and y and graph the two resulting equations. The combined graphs will show a 90° rotation of the graph of the original equation.

b. Set the graphing utility to *parametric* mode and graph the equations

$$x = -\sqrt{2t^3 - 4t + 2}$$
$$y = t$$

and

$$x = \sqrt{2t^3 - 4t + 2}$$
$$y = t.$$

From either of these two approaches, can you decide whether the graph has a tangent line at the point $(0, 1)$? Explain your reasoning.

Implicit and Explicit Functions

Up to this point in the text, most functions have been expressed in **explicit form.** For example, in the equation

$$y = 3x^2 - 5 \qquad \text{Explicit form}$$

the variable y is explicitly written as a function of x. Some functions, however, are only implied by an equation. For instance, the function $y = 1/x$ is defined **implicitly** by the equation $xy = 1$. Suppose you were asked to find dy/dx for this equation. You could begin by writing y explicitly as a function of x and then differentiating.

Implicit Form	*Explicit Form*	*Derivative*
$xy = 1$	$y = \dfrac{1}{x} = x^{-1}$	$\dfrac{dy}{dx} = -x^{-2} = -\dfrac{1}{x^2}$

This strategy works whenever you can solve for the function explicitly. You cannot, however, use this procedure when you are unable to solve for y as a function of x. For instance, how would you find dy/dx for the equation

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

where it is very difficult to express y as a function of x explicitly? To do this, you can use **implicit differentiation.**

To understand how to find dy/dx implicitly, you must realize that the differentiation is taking place *with respect to x*. This means that when you differentiate terms involving x alone, you can differentiate as usual. However, when you differentiate terms involving y, you must apply the Chain Rule, because you are assuming that y is defined implicitly as a differentiable function of x.

Video

EXAMPLE 1 Differentiating with Respect to *x*

a. $\dfrac{d}{dx}[x^3] = 3x^2$ Variables agree: use Simple Power Rule.

Variables agree

b. $\dfrac{d}{dx}[\underbrace{y^3}_{u^n}] = \underbrace{3y^2}_{nu^{n-1}}\underbrace{\dfrac{dy}{dx}}_{u'}$ Variables disagree: use Chain Rule.

Variables disagree

c. $\dfrac{d}{dx}[x + 3y] = 1 + 3\dfrac{dy}{dx}$ Chain Rule: $\dfrac{d}{dx}[3y] = 3y'$

d. $\dfrac{d}{dx}[xy^2] = x\dfrac{d}{dx}[y^2] + y^2\dfrac{d}{dx}[x]$ Product Rule

$= x\left(2y\dfrac{dy}{dx}\right) + y^2(1)$ Chain Rule

$= 2xy\dfrac{dy}{dx} + y^2$ Simplify.

Try It Exploration A

Implicit Differentiation

Guidelines for Implicit Differentiation

1. Differentiate both sides of the equation *with respect to x*.
2. Collect all terms involving dy/dx on the left side of the equation and move all other terms to the right side of the equation.
3. Factor dy/dx out of the left side of the equation.
4. Solve for dy/dx.

EXAMPLE 2 Implicit Differentiation

Find dy/dx given that $y^3 + y^2 - 5y - x^2 = -4$.

Solution

NOTE In Example 2, note that implicit differentiation can produce an expression for dy/dx that contains both x and y.

1. Differentiate both sides of the equation with respect to x.

$$\frac{d}{dx}[y^3 + y^2 - 5y - x^2] = \frac{d}{dx}[-4]$$

$$\frac{d}{dx}[y^3] + \frac{d}{dx}[y^2] - \frac{d}{dx}[5y] - \frac{d}{dx}[x^2] = \frac{d}{dx}[-4]$$

$$3y^2\frac{dy}{dx} + 2y\frac{dy}{dx} - 5\frac{dy}{dx} - 2x = 0$$

2. Collect the dy/dx terms on the left side of the equation and move all other terms to the right side of the equation.

$$3y^2\frac{dy}{dx} + 2y\frac{dy}{dx} - 5\frac{dy}{dx} = 2x$$

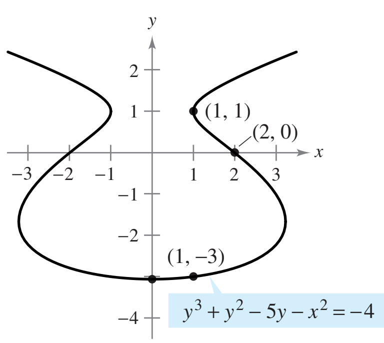
3. Factor dy/dx out of the left side of the equation.

$$\frac{dy}{dx}(3y^2 + 2y - 5) = 2x$$

4. Solve for dy/dx by dividing by $(3y^2 + 2y - 5)$.

$$\frac{dy}{dx} = \frac{2x}{3y^2 + 2y - 5}$$

Try It Exploration A Video Video



Point on Graph	*Slope of Graph*
(2, 0)	$-\frac{4}{5}$
$(1, -3)$	$\frac{1}{8}$
$x = 0$	0
(1, 1)	Undefined

The implicit equation

$$y^3 + y^2 - 5y - x^2 = -4$$

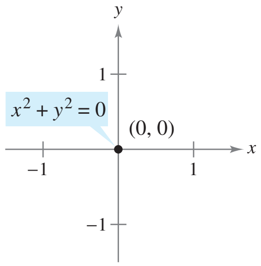
has the derivative

$$\frac{dy}{dx} = \frac{2x}{3y^2 + 2y - 5}.$$

Figure 2.27

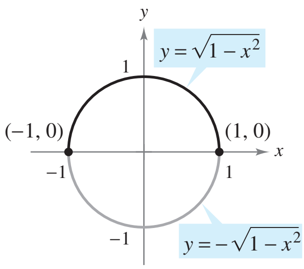
To see how you can use an *implicit derivative*, consider the graph shown in Figure 2.27. From the graph, you can see that y is not a function of x. Even so, the derivative found in Example 2 gives a formula for the slope of the tangent line at a point on this graph. The slopes at several points on the graph are shown below the graph.

TECHNOLOGY With most graphing utilities, it is easy to graph an equation that explicitly represents y as a function of x. Graphing other equations, however, can require some ingenuity. For instance, to graph the equation given in Example 2, use a graphing utility, set in *parametric* mode, to graph the parametric representations $x = \sqrt{t^3 + t^2 - 5t + 4}$, $y = t$, and $x = -\sqrt{t^3 + t^2 - 5t + 4}$, $y = t$, for $-5 \le t \le 5$. How does the result compare with the graph shown in Figure 2.27?



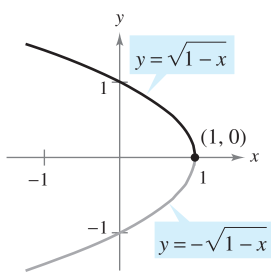
(a)

Editable Graph



(b)

Editable Graph



(c)

Editable Graph

Some graph segments can be represented by differentiable functions.
Figure 2.28

It is meaningless to solve for dy/dx in an equation that has no solution points. (For example, $x^2 + y^2 = -4$ has no solution points.) If, however, a segment of a graph can be represented by a differentiable function, dy/dx will have meaning as the slope at each point on the segment. Recall that a function is not differentiable at (a) points with vertical tangents and (b) points at which the function is not continuous.

EXAMPLE 3 Representing a Graph by Differentiable Functions

If possible, represent y as a differentiable function of x.

a. $x^2 + y^2 = 0$ **b.** $x^2 + y^2 = 1$ **c.** $x + y^2 = 1$

Solution

a. The graph of this equation is a single point. So, it does not define y as a differentiable function of x. See Figure 2.28(a).

b. The graph of this equation is the unit circle, centered at $(0, 0)$. The upper semicircle is given by the differentiable function

$$y = \sqrt{1 - x^2}, \quad -1 < x < 1$$

and the lower semicircle is given by the differentiable function

$$y = -\sqrt{1 - x^2}, \quad -1 < x < 1.$$

At the points $(-1, 0)$ and $(1, 0)$, the slope of the graph is undefined. See Figure 2.28(b).

c. The upper half of this parabola is given by the differentiable function

$$y = \sqrt{1 - x}, \quad x < 1$$

and the lower half of this parabola is given by the differentiable function

$$y = -\sqrt{1 - x}, \quad x < 1.$$

At the point $(1, 0)$, the slope of the graph is undefined. See Figure 2.28(c).

Try It | Exploration A | Exploration B

EXAMPLE 4 Finding the Slope of a Graph Implicitly

Determine the slope of the tangent line to the graph of

$$x^2 + 4y^2 = 4$$

at the point $\left(\sqrt{2}, -1/\sqrt{2}\right)$. See Figure 2.29.

Solution

$$x^2 + 4y^2 = 4 \qquad \text{Write original equation.}$$

$$2x + 8y\frac{dy}{dx} = 0 \qquad \text{Differentiate with respect to } x.$$

$$\frac{dy}{dx} = \frac{-2x}{8y} = \frac{-x}{4y} \qquad \text{Solve for } \frac{dy}{dx}.$$

So, at $\left(\sqrt{2}, -1/\sqrt{2}\right)$, the slope is

$$\frac{dy}{dx} = \frac{-\sqrt{2}}{-4/\sqrt{2}} = \frac{1}{2}. \qquad \text{Evaluate } \frac{dy}{dx} \text{ when } x = \sqrt{2} \text{ and } y = -\frac{1}{\sqrt{2}}.$$

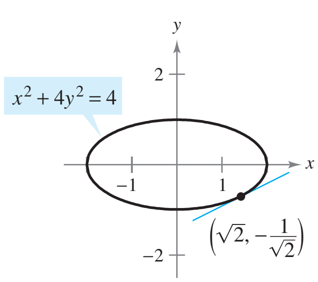


Figure 2.29

Editable Graph

Try It | Exploration A | Exploration B | Open Exploration

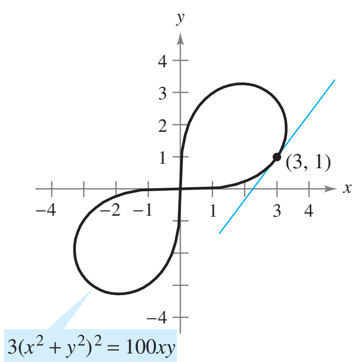
NOTE To see the benefit of implicit differentiation, try doing Example 4 using the explicit function $y = -\frac{1}{2}\sqrt{4 - x^2}$.

EXAMPLE 5 Finding the Slope of a Graph Implicitly

Determine the slope of the graph of $3(x^2 + y^2)^2 = 100xy$ at the point $(3, 1)$.

Solution

$$\frac{d}{dx}[3(x^2 + y^2)^2] = \frac{d}{dx}[100xy]$$
$$3(2)(x^2 + y^2)\left(2x + 2y\frac{dy}{dx}\right) = 100\left[x\frac{dy}{dx} + y(1)\right]$$
$$12y(x^2 + y^2)\frac{dy}{dx} - 100x\frac{dy}{dx} = 100y - 12x(x^2 + y^2)$$
$$[12y(x^2 + y^2) - 100x]\frac{dy}{dx} = 100y - 12x(x^2 + y^2)$$
$$\frac{dy}{dx} = \frac{100y - 12x(x^2 + y^2)}{-100x + 12y(x^2 + y^2)}$$
$$= \frac{25y - 3x(x^2 + y^2)}{-25x + 3y(x^2 + y^2)}$$



Lemniscate
Figure 2.30

At the point $(3, 1)$, the slope of the graph is

$$\frac{dy}{dx} = \frac{25(1) - 3(3)(3^2 + 1^2)}{-25(3) + 3(1)(3^2 + 1^2)} = \frac{25 - 90}{-75 + 30} = \frac{-65}{-45} = \frac{13}{9}$$

as shown in Figure 2.30. This graph is called a **lemniscate.**

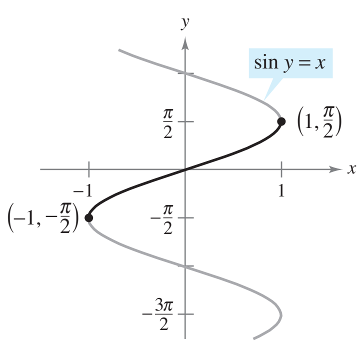
Try It | Exploration A | Exploration B

EXAMPLE 6 Determining a Differentiable Function

Find dy/dx implicitly for the equation $\sin y = x$. Then find the largest interval of the form $-a < y < a$ on which y is a differentiable function of x (see Figure 2.31).

Solution

$$\frac{d}{dx}[\sin y] = \frac{d}{dx}[x]$$
$$\cos y\frac{dy}{dx} = 1$$
$$\frac{dy}{dx} = \frac{1}{\cos y}$$



The derivative is $\dfrac{dy}{dx} = \dfrac{1}{\sqrt{1 - x^2}}$.
Figure 2.31

Editable Graph

The largest interval about the origin for which y is a differentiable function of x is $-\pi/2 < y < \pi/2$. To see this, note that $\cos y$ is positive for all y in this interval and is 0 at the endpoints. If you restrict y to the interval $-\pi/2 < y < \pi/2$, you should be able to write dy/dx explicitly as a function of x. To do this, you can use

$$\cos y = \sqrt{1 - \sin^2 y}$$
$$= \sqrt{1 - x^2}, \quad -\frac{\pi}{2} < y < \frac{\pi}{2}$$

and conclude that

$$\frac{dy}{dx} = \frac{1}{\sqrt{1 - x^2}}.$$

Try It | Exploration A

Isaac Barrow (1630–1677)

The graph in Figure 2.32 is called the **kappa curve** because it resembles the Greek letter kappa, κ. The general solution for the tangent line to this curve was discovered by the English mathematician Isaac Barrow. Newton was Barrow's student, and they corresponded frequently regarding their work in the early development of calculus.

MathBio

With implicit differentiation, the form of the derivative often can be simplified (as in Example 6) by an appropriate use of the *original* equation. A similar technique can be used to find and simplify higher-order derivatives obtained implicitly.

EXAMPLE 7 Finding the Second Derivative Implicitly

Given $x^2 + y^2 = 25$, find $\dfrac{d^2y}{dx^2}$.

Solution Differentiating each term with respect to x produces

$$2x + 2y\frac{dy}{dx} = 0$$

$$2y\frac{dy}{dx} = -2x$$

$$\frac{dy}{dx} = \frac{-2x}{2y} = -\frac{x}{y}.$$

Differentiating a second time with respect to x yields

$$\frac{d^2y}{dx^2} = -\frac{(y)(1) - (x)(dy/dx)}{y^2} \qquad \text{Quotient Rule}$$

$$= -\frac{y - (x)(-x/y)}{y^2} \qquad \text{Substitute } -x/y \text{ for } \frac{dy}{dx}.$$

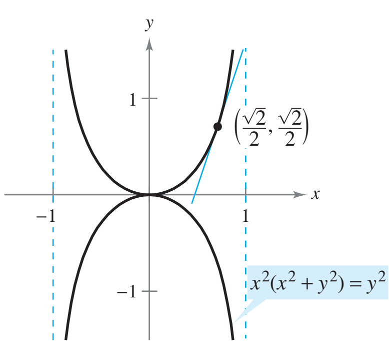
$$= -\frac{y^2 + x^2}{y^3} \qquad \text{Simplify.}$$

$$= -\frac{25}{y^3}. \qquad \text{Substitute 25 for } x^2 + y^2.$$

Try It Exploration A

EXAMPLE 8 Finding a Tangent Line to a Graph

Find the tangent line to the graph given by $x^2(x^2 + y^2) = y^2$ at the point $(\sqrt{2}/2, \sqrt{2}/2)$, as shown in Figure 2.32.



The kappa curve
Figure 2.32

Solution By rewriting and differentiating implicitly, you obtain

$$x^4 + x^2y^2 - y^2 = 0$$

$$4x^3 + x^2\left(2y\frac{dy}{dx}\right) + 2xy^2 - 2y\frac{dy}{dx} = 0$$

$$2y(x^2 - 1)\frac{dy}{dx} = -2x(2x^2 + y^2)$$

$$\frac{dy}{dx} = \frac{x(2x^2 + y^2)}{y(1 - x^2)}.$$

At the point $(\sqrt{2}/2, \sqrt{2}/2)$, the slope is

$$\frac{dy}{dx} = \frac{(\sqrt{2}/2)[2(1/2) + (1/2)]}{(\sqrt{2}/2)[1 - (1/2)]} = \frac{3/2}{1/2} = 3$$

and the equation of the tangent line at this point is

$$y - \frac{\sqrt{2}}{2} = 3\left(x - \frac{\sqrt{2}}{2}\right)$$

$$y = 3x - \sqrt{2}.$$

Try It Exploration A

Section 2.6 Related Rates

- **Find a related rate.**
- **Use related rates to solve real-life problems.**

Finding Related Rates

You have seen how the Chain Rule can be used to find dy/dx implicitly. Another important use of the Chain Rule is to find the rates of change of two or more related variables that are changing with respect to *time*.

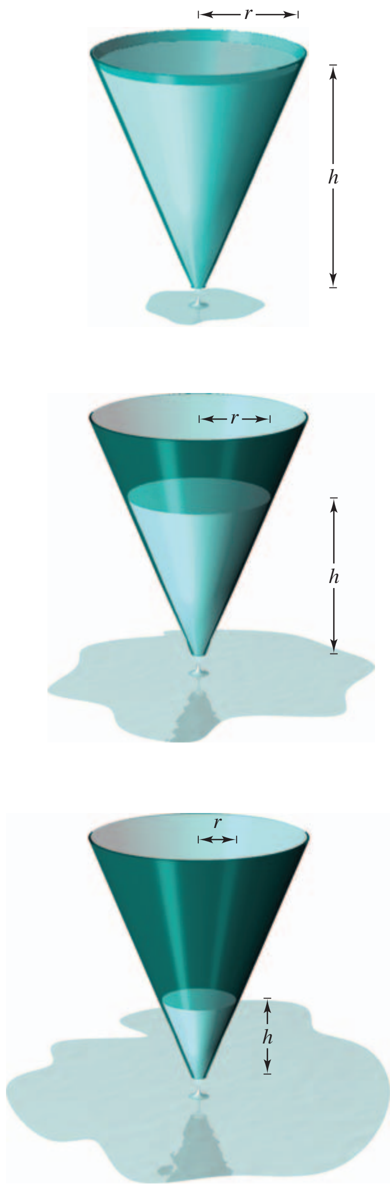
For example, when water is drained out of a conical tank (see Figure 2.33), the volume V, the radius r, and the height h of the water level are all functions of time t. Knowing that these variables are related by the equation

$$V = \frac{\pi}{3}r^2h \qquad \text{Original equation}$$

you can differentiate implicitly with respect to t to obtain the **related-rate** equation

$$\begin{aligned}
\frac{d}{dt}(V) &= \frac{d}{dt}\left(\frac{\pi}{3}r^2h\right) \\
\frac{dV}{dt} &= \frac{\pi}{3}\left[r^2\frac{dh}{dt} + h\left(2r\frac{dr}{dt}\right)\right] && \text{Differentiate with respect to } t. \\
&= \frac{\pi}{3}\left(r^2\frac{dh}{dt} + 2rh\frac{dr}{dt}\right).
\end{aligned}$$

From this equation you can see that the rate of change of V is related to the rates of change of both h and r.



Volume is related to radius and height.
Figure 2.33

Animation

> **EXPLORATION**
>
> ***Finding a Related Rate*** In the conical tank shown in Figure 2.33, suppose that the height is changing at a rate of -0.2 foot per minute and the radius is changing at a rate of -0.1 foot per minute. What is the rate of change in the volume when the radius is $r = 1$ foot and the height is $h = 2$ feet? Does the rate of change in the volume depend on the values of r and h? Explain.

EXAMPLE 1 Two Rates That Are Related

Suppose x and y are both differentiable functions of t and are related by the equation $y = x^2 + 3$. Find dy/dt when $x = 1$, given that $dx/dt = 2$ when $x = 1$.

Solution Using the Chain Rule, you can differentiate both sides of the equation *with respect to t*.

$$\begin{aligned}
y &= x^2 + 3 && \text{Write original equation.} \\
\frac{d}{dt}[y] &= \frac{d}{dt}[x^2 + 3] && \text{Differentiate with respect to } t. \\
\frac{dy}{dt} &= 2x\frac{dx}{dt} && \text{Chain Rule}
\end{aligned}$$

When $x = 1$ and $dx/dt = 2$, you have

$$\frac{dy}{dt} = 2(1)(2) = 4.$$

Try It Exploration A

FOR FURTHER INFORMATION To learn more about the history of related-rate problems, see the article "The Lengthening Shadow: The Story of Related Rates" by Bill Austin, Don Barry, and David Berman in *Mathematics Magazine*.

MathArticle

Problem Solving with Related Rates

In Example 1, you were *given* an equation that related the variables x and y and were asked to find the rate of change of y when $x = 1$.

Equation: $y = x^2 + 3$

Given rate: $\dfrac{dx}{dt} = 2 \quad \text{when} \quad x = 1$

Find: $\dfrac{dy}{dt} \quad \text{when} \quad x = 1$

In each of the remaining examples in this section, you must *create* a mathematical model from a verbal description.

EXAMPLE 2 Ripples in a Pond

A pebble is dropped into a calm pond, causing ripples in the form of concentric circles, as shown in Figure 2.34. The radius r of the outer ripple is increasing at a constant rate of 1 foot per second. When the radius is 4 feet, at what rate is the total area A of the disturbed water changing?

Total area increases as the outer radius increases.
Figure 2.34

Solution The variables r and A are related by $A = \pi r^2$. The rate of change of the radius r is $dr/dt = 1$.

Equation: $A = \pi r^2$

Given rate: $\dfrac{dr}{dt} = 1$

Find: $\dfrac{dA}{dt} \quad \text{when} \quad r = 4$

With this information, you can proceed as in Example 1.

$$\frac{d}{dt}[A] = \frac{d}{dt}[\pi r^2] \qquad \text{Differentiate with respect to } t.$$

$$\frac{dA}{dt} = 2\pi r \frac{dr}{dt} \qquad \text{Chain Rule}$$

$$\frac{dA}{dt} = 2\pi(4)(1) = 8\pi \qquad \text{Substitute 4 for } r \text{ and 1 for } dr/dt.$$

When the radius is 4 feet, the area is changing at a rate of 8π square feet per second.

Try It | Exploration A | Video | Video

NOTE When using these guidelines, be sure you perform Step 3 before Step 4. Substituting the known values of the variables before differentiating will produce an inappropriate derivative.

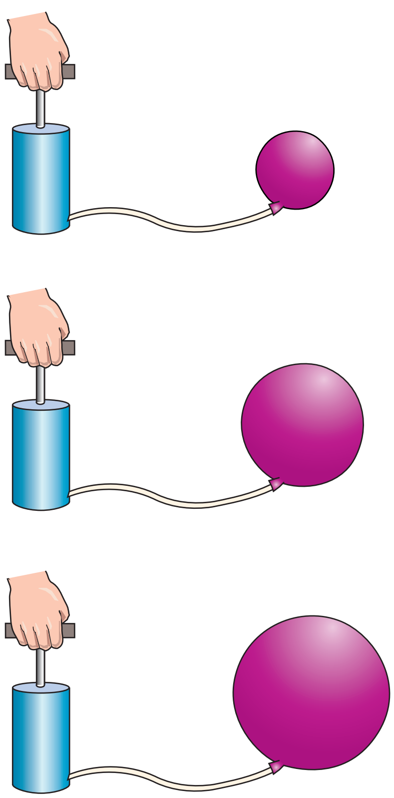
Guidelines For Solving Related-Rate Problems

1. Identify all *given* quantities and quantities *to be determined*. Make a sketch and label the quantities.
2. Write an equation involving the variables whose rates of change either are given or are to be determined.
3. Using the Chain Rule, implicitly differentiate both sides of the equation *with respect to time t*.
4. *After* completing Step 3, substitute into the resulting equation all known values for the variables and their rates of change. Then solve for the required rate of change.

The table below lists examples of mathematical models involving rates of change. For instance, the rate of change in the first example is the velocity of a car.

Verbal Statement	Mathematical Model
The velocity of a car after traveling for 1 hour is 50 miles per hour.	$x =$ distance traveled $\dfrac{dx}{dt} = 50$ when $t = 1$
Water is being pumped into a swimming pool at a rate of 10 cubic meters per hour.	$V =$ volume of water in pool $\dfrac{dV}{dt} = 10 \text{ m}^3/\text{hr}$
A gear is revolving at a rate of 25 revolutions per minute (1 revolution $= 2\pi$ rad).	$\theta =$ angle of revolution $\dfrac{d\theta}{dt} = 25(2\pi)$ rad/min

EXAMPLE 3 An Inflating Balloon



Inflating a balloon
Figure 2.35

Animation

Air is being pumped into a spherical balloon (see Figure 2.35) at a rate of 4.5 cubic feet per minute. Find the rate of change of the radius when the radius is 2 feet.

Solution Let V be the volume of the balloon and let r be its radius. Because the volume is increasing at a rate of 4.5 cubic feet per minute, you know that at time t the rate of change of the volume is $dV/dt = \frac{9}{2}$. So, the problem can be stated as shown.

Given rate: $\dfrac{dV}{dt} = \dfrac{9}{2}$ (constant rate)

Find: $\dfrac{dr}{dt}$ when $r = 2$

To find the rate of change of the radius, you must find an equation that relates the radius r to the volume V.

Equation: $V = \dfrac{4}{3}\pi r^3$ Volume of a sphere

Differentiating both sides of the equation with respect to t produces

$$\frac{dV}{dt} = 4\pi r^2 \frac{dr}{dt} \qquad \text{Differentiate with respect to } t.$$

$$\frac{dr}{dt} = \frac{1}{4\pi r^2}\left(\frac{dV}{dt}\right). \qquad \text{Solve for } dr/dt.$$

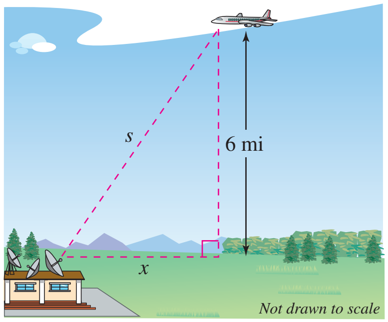
Finally, when $r = 2$, the rate of change of the radius is

$$\frac{dr}{dt} = \frac{1}{16\pi}\left(\frac{9}{2}\right) \approx 0.09 \text{ foot per minute.}$$

Try It | Exploration A | Video

In Example 3, note that the volume is increasing at a *constant* rate but the radius is increasing at a *variable* rate. Just because two rates are related does not mean that they are proportional. In this particular case, the radius is growing more and more slowly as t increases. Do you see why?

EXAMPLE 4 The Speed of an Airplane Tracked by Radar



An airplane is flying at an altitude of 6 miles, s miles from the station.
Figure 2.36

An airplane is flying on a flight path that will take it directly over a radar tracking station, as shown in Figure 2.36. If s is decreasing at a rate of 400 miles per hour when $s = 10$ miles, what is the speed of the plane?

Solution Let x be the horizontal distance from the station, as shown in Figure 2.36. Notice that when $s = 10$, $x = \sqrt{10^2 - 36} = 8$.

Given rate: $ds/dt = -400$ when $s = 10$

Find: dx/dt when $s = 10$ and $x = 8$

You can find the velocity of the plane as shown.

Equation: $x^2 + 6^2 = s^2$ Pythagorean Theorem

$$2x\frac{dx}{dt} = 2s\frac{ds}{dt}$$ Differentiate with respect to t.

$$\frac{dx}{dt} = \frac{s}{x}\left(\frac{ds}{dt}\right)$$ Solve for dx/dt.

$$\frac{dx}{dt} = \frac{10}{8}(-400)$$ Substitute for s, x, and ds/dt.

$$= -500 \text{ miles per hour}$$ Simplify.

Because the velocity is -500 miles per hour, the *speed* is 500 miles per hour.

Try It | Exploration A | Open Exploration

EXAMPLE 5 A Changing Angle of Elevation

Find the rate of change in the angle of elevation of the camera shown in Figure 2.37 at 10 seconds after lift-off.

Solution Let θ be the angle of elevation, as shown in Figure 2.37. When $t = 10$, the height s of the rocket is $s = 50t^2 = 50(10)^2 = 5000$ feet.

Given rate: $ds/dt = 100t$ = velocity of rocket

Find: $d\theta/dt$ when $t = 10$ and $s = 5000$

Using Figure 2.37, you can relate s and θ by the equation $\tan\theta = s/2000$.

Equation: $\tan\theta = \dfrac{s}{2000}$ See Figure 2.37.

$$(\sec^2\theta)\frac{d\theta}{dt} = \frac{1}{2000}\left(\frac{ds}{dt}\right)$$ Differentiate with respect to t.

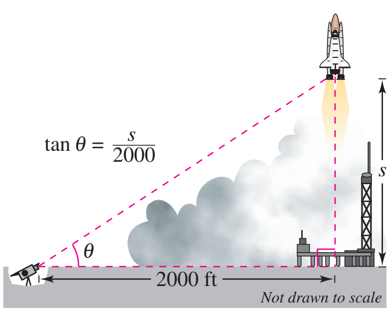
$$\frac{d\theta}{dt} = \cos^2\theta\,\frac{100t}{2000}$$ Substitute $100t$ for ds/dt.

$$= \left(\frac{2000}{\sqrt{s^2 + 2000^2}}\right)^2 \frac{100t}{2000}$$ $\cos\theta = 2000/\sqrt{s^2 + 2000^2}$

When $t = 10$ and $s = 5000$, you have

$$\frac{d\theta}{dt} = \frac{2000(100)(10)}{5000^2 + 2000^2} = \frac{2}{29} \text{ radian per second.}$$

So, when $t = 10$, θ is changing at a rate of $\frac{2}{29}$ radian per second.



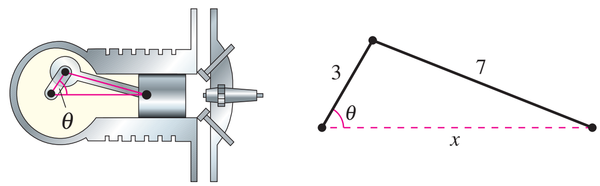
A television camera at ground level is filming the lift-off of a space shuttle that is rising vertically according to the position equation $s = 50t^2$, where s is measured in feet and t is measured in seconds. The camera is 2000 feet from the launch pad.
Figure 2.37

Animation

Try It | Exploration A

EXAMPLE 6 The Velocity of a Piston

In the engine shown in Figure 2.38, a 7-inch connecting rod is fastened to a crank of radius 3 inches. The crankshaft rotates counterclockwise at a constant rate of 200 revolutions per minute. Find the velocity of the piston when $\theta = \pi/3$.



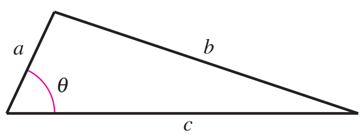
The velocity of a piston is related to the angle of the crankshaft.
Figure 2.38

Animation

Solution Label the distances as shown in Figure 2.38. Because a complete revolution corresponds to 2π radians, it follows that $d\theta/dt = 200(2\pi) = 400\pi$ radians per minute.

Given rate: $\dfrac{d\theta}{dt} = 400\pi$ (constant rate)

Find: $\dfrac{dx}{dt}$ when $\theta = \dfrac{\pi}{3}$



Law of Cosines:
$b^2 = a^2 + c^2 - 2ac\cos\theta$
Figure 2.39

You can use the Law of Cosines (Figure 2.39) to find an equation that relates x and θ.

Equation:

$$7^2 = 3^2 + x^2 - 2(3)(x)\cos\theta$$

$$0 = 2x\frac{dx}{dt} - 6\left(-x\sin\theta\frac{d\theta}{dt} + \cos\theta\frac{dx}{dt}\right)$$

$$(6\cos\theta - 2x)\frac{dx}{dt} = 6x\sin\theta\frac{d\theta}{dt}$$

$$\frac{dx}{dt} = \frac{6x\sin\theta}{6\cos\theta - 2x}\left(\frac{d\theta}{dt}\right)$$

When $\theta = \pi/3$, you can solve for x as shown.

$$7^2 = 3^2 + x^2 - 2(3)(x)\cos\frac{\pi}{3}$$

$$49 = 9 + x^2 - 6x\left(\frac{1}{2}\right)$$

$$0 = x^2 - 3x - 40$$

$$0 = (x - 8)(x + 5)$$

$$x = 8$$ Choose positive solution.

So, when $x = 8$ and $\theta = \pi/3$, the velocity of the piston is

$$\frac{dx}{dt} = \frac{6(8)\left(\sqrt{3}/2\right)}{6(1/2) - 16}(400\pi)$$

$$= \frac{9600\pi\sqrt{3}}{-13}$$

$$\approx -4018 \text{ inches per minute.}$$

Try It Exploration A

NOTE Note that the velocity in Example 6 is negative because x represents a distance that is decreasing.

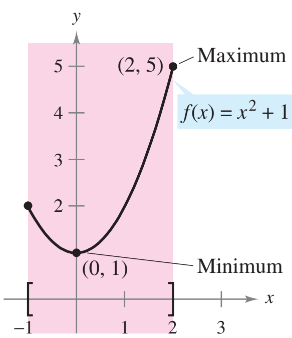
Section 3.1 Extrema on an Interval

- **Understand the definition of extrema of a function on an interval.**
- **Understand the definition of relative extrema of a function on an open interval.**
- **Find extrema on a closed interval.**

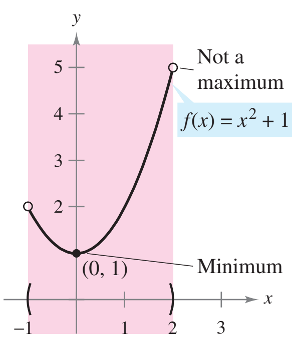
Video

Extrema of a Function

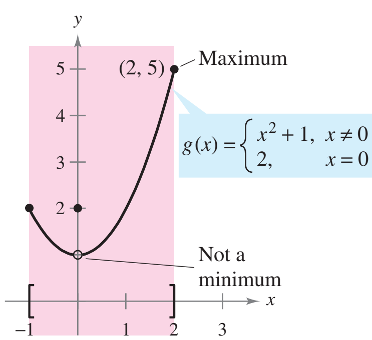
In calculus, much effort is devoted to determining the behavior of a function f on an interval I. Does f have a maximum value on I? Does it have a minimum value? Where is the function increasing? Where is it decreasing? In this chapter you will learn how derivatives can be used to answer these questions. You will also see why these questions are important in real-life applications.



(a) f is continuous, $[-1, 2]$ is closed.



(b) f is continuous, $(-1, 2)$ is open.



(c) g is not continuous, $[-1, 2]$ is closed.

Extrema can occur at interior points or endpoints of an interval. Extrema that occur at the endpoints are called **endpoint extrema.**

Figure 3.1

> **Definition of Extrema**
>
> Let f be defined on an interval I containing c.
>
> **1.** $f(c)$ is the **minimum of f on I** if $f(c) \leq f(x)$ for all x in I.
>
> **2.** $f(c)$ is the **maximum of f on I** if $f(c) \geq f(x)$ for all x in I.
>
> The minimum and maximum of a function on an interval are the **extreme values,** or **extrema** (the singular form of extrema is extremum), of the function on the interval. The minimum and maximum of a function on an interval are also called the **absolute minimum** and **absolute maximum** on the interval.

A function need not have a minimum or a maximum on an interval. For instance, in Figure 3.1(a) and (b), you can see that the function $f(x) = x^2 + 1$ has both a minimum and a maximum on the closed interval $[-1, 2]$, but does not have a maximum on the open interval $(-1, 2)$. Moreover, in Figure 3.1(c), you can see that continuity (or the lack of it) can affect the existence of an extremum on the interval. This suggests the theorem below. (Although the Extreme Value Theorem is intuitively plausible, a proof of this theorem is not within the scope of this text.)

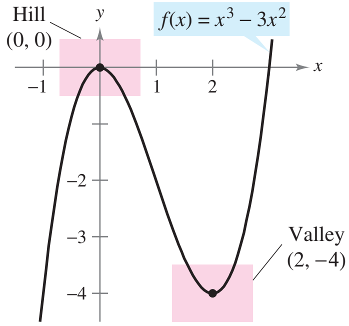
> **THEOREM 3.1 The Extreme Value Theorem**
>
> If f is continuous on a closed interval $[a, b]$, then f has both a minimum and a maximum on the interval.

EXPLORATION

Finding Minimum and Maximum Values The Extreme Value Theorem (like the Intermediate Value Theorem) is an *existence theorem* because it tells of the existence of minimum and maximum values but does not show how to find these values. Use the extreme-value capability of a graphing utility to find the minimum and maximum values of each of the following functions. In each case, do you think the x-values are exact or approximate? Explain your reasoning.

a. $f(x) = x^2 - 4x + 5$ on the closed interval $[-1, 3]$

b. $f(x) = x^3 - 2x^2 - 3x - 2$ on the closed interval $[-1, 3]$



f has a relative maximum at $(0, 0)$ and a relative minimum at $(2, -4)$.
Figure 3.2

Relative Extrema and Critical Numbers

In Figure 3.2, the graph of $f(x) = x^3 - 3x^2$ has a **relative maximum** at the point $(0, 0)$ and a **relative minimum** at the point $(2, -4)$. Informally, you can think of a relative maximum as occurring on a "hill" on the graph, and a relative minimum as occurring in a "valley" on the graph. Such a hill and valley can occur in two ways. If the hill (or valley) is smooth and rounded, the graph has a horizontal tangent line at the high point (or low point). If the hill (or valley) is sharp and peaked, the graph represents a function that is not differentiable at the high point (or low point).

Definition of Relative Extrema

1. If there is an open interval containing c on which $f(c)$ is a maximum, then $f(c)$ is called a **relative maximum** of f, or you can say that f has a **relative maximum at $(c, f(c))$.**
2. If there is an open interval containing c on which $f(c)$ is a minimum, then $f(c)$ is called a **relative minimum** of f, or you can say that f has a **relative minimum at $(c, f(c))$.**

The plural of relative maximum is relative maxima, and the plural of relative minimum is relative minima.

Example 1 examines the derivatives of functions at *given* relative extrema. (Much more is said about *finding* the relative extrema of a function in Section 3.3.)

EXAMPLE 1 The Value of the Derivative at Relative Extrema

Find the value of the derivative at each of the relative extrema shown in Figure 3.3.

Solution

a. The derivative of $f(x) = \dfrac{9(x^2 - 3)}{x^3}$ is

$$f'(x) = \frac{x^3(18x) - (9)(x^2 - 3)(3x^2)}{(x^3)^2} \qquad \text{Differentiate using Quotient Rule.}$$

$$= \frac{9(9 - x^2)}{x^4}. \qquad \text{Simplify.}$$

At the point $(3, 2)$, the value of the derivative is $f'(3) = 0$ [see Figure 3.3(a)].

b. At $x = 0$, the derivative of $f(x) = |x|$ *does not exist* because the following one-sided limits differ [see Figure 3.3(b)].

$$\lim_{x \to 0^-} \frac{f(x) - f(0)}{x - 0} = \lim_{x \to 0^-} \frac{|x|}{x} = -1 \qquad \text{Limit from the left}$$

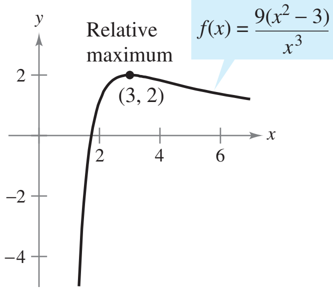
$$\lim_{x \to 0^+} \frac{f(x) - f(0)}{x - 0} = \lim_{x \to 0^+} \frac{|x|}{x} = 1 \qquad \text{Limit from the right}$$

c. The derivative of $f(x) = \sin x$ is

$$f'(x) = \cos x.$$

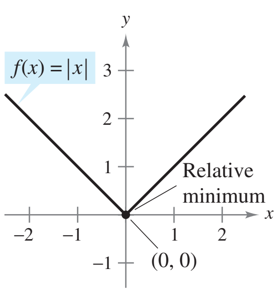
At the point $(\pi/2, 1)$, the value of the derivative is $f'(\pi/2) = \cos(\pi/2) = 0$. At the point $(3\pi/2, -1)$, the value of the derivative is $f'(3\pi/2) = \cos(3\pi/2) = 0$ [see Figure 3.3(c)].

Try It | Exploration A | Exploration B



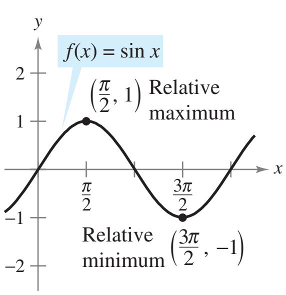
(a) $f'(3) = 0$

Editable Graph



(b) $f'(0)$ does not exist.

Editable Graph



(c) $f'\left(\dfrac{\pi}{2}\right) = 0$; $f'\left(\dfrac{3\pi}{2}\right) = 0$

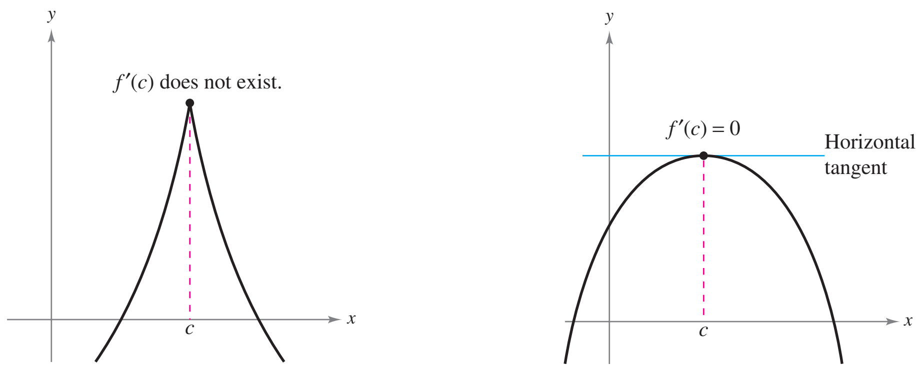
Editable Graph

Figure 3.3

Note in Example 1 that at the relative extrema, the derivative is either zero or does not exist. The x-values at these special points are called **critical numbers.** Figure 3.4 illustrates the two types of critical numbers.

Definition of a Critical Number

Let f be defined at c. If $f'(c) = 0$ or if f is not differentiable at c, then c is a **critical number** of f.



c is a critical number of *f*.
Figure 3.4

THEOREM 3.2 Relative Extrema Occur Only at Critical Numbers

If f has a relative minimum or relative maximum at $x = c$, then c is a critical number of f.

Proof

Case 1: If f is *not* differentiable at $x = c$, then, by definition, c is a critical number of f and the theorem is valid.

Case 2: If f is differentiable at $x = c$, then $f'(c)$ must be positive, negative, or 0. Suppose $f'(c)$ is positive. Then

$$f'(c) = \lim_{x \to c} \frac{f(x) - f(c)}{x - c} > 0$$

which implies that there exists an interval (a, b) containing c such that

$$\frac{f(x) - f(c)}{x - c} > 0, \text{ for all } x \neq c \text{ in } (a, b).$$

[See Exercise 72(b), Section 1.2.]

Because this quotient is positive, the signs of the denominator and numerator must agree. This produces the following inequalities for x-values in the interval (a, b).

Left of c: $x < c$ and $f(x) < f(c)$ $\Rightarrow$ $f(c)$ is not a relative minimum

Right of c: $x > c$ and $f(x) > f(c)$ $\Rightarrow$ $f(c)$ is not a relative maximum

So, the assumption that $f'(c) > 0$ contradicts the hypothesis that $f(c)$ is a relative extremum. Assuming that $f'(c) < 0$ produces a similar contradiction, you are left with only one possibility—namely, $f'(c) = 0$. So, by definition, c is a critical number of f and the theorem is valid.

PIERRE DE FERMAT (1601–1665)

For Fermat, who was trained as a lawyer, mathematics was more of a hobby than a profession. Nevertheless, Fermat made many contributions to analytic geometry, number theory, calculus, and probability. In letters to friends, he wrote of many of the fundamental ideas of calculus, long before Newton or Leibniz. For instance, the theorem at the right is sometimes attributed to Fermat.

Finding Extrema on a Closed Interval

Theorem 3.2 states that the relative extrema of a function can occur *only* at the critical numbers of the function. Knowing this, you can use the following guidelines to find extrema on a closed interval.

> **Guidelines for Finding Extrema on a Closed Interval**
>
> To find the extrema of a continuous function f on a closed interval $[a, b]$, use the following steps.
>
> **1.** Find the critical numbers of f in (a, b).
>
> **2.** Evaluate f at each critical number in (a, b).
>
> **3.** Evaluate f at each endpoint of $[a, b]$.
>
> **4.** The least of these values is the minimum. The greatest is the maximum.

Technology

The next three examples show how to apply these guidelines. Be sure you see that finding the critical numbers of the function is only part of the procedure. Evaluating the function at the critical numbers *and* the endpoints is the other part.

EXAMPLE 2 Finding Extrema on a Closed Interval

Find the extrema of $f(x) = 3x^4 - 4x^3$ on the interval $[-1, 2]$.

Solution Begin by differentiating the function.

$$f(x) = 3x^4 - 4x^3 \qquad \text{Write original function.}$$

$$f'(x) = 12x^3 - 12x^2 \qquad \text{Differentiate.}$$

To find the critical numbers of f, you must find all x-values for which $f'(x) = 0$ and all x-values for which $f'(x)$ does not exist.

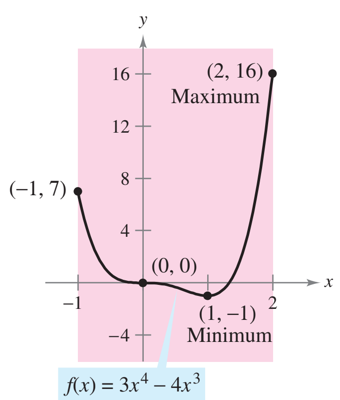
$$f'(x) = 12x^3 - 12x^2 = 0 \qquad \text{Set } f'(x) \text{ equal to 0.}$$

$$12x^2(x - 1) = 0 \qquad \text{Factor.}$$

$$x = 0, 1 \qquad \text{Critical numbers}$$

Because f' is defined for all x, you can conclude that these are the only critical numbers of f. By evaluating f at these two critical numbers and at the endpoints of $[-1, 2]$, you can determine that the maximum is $f(2) = 16$ and the minimum is $f(1) = -1$, as shown in the table. The graph of f is shown in Figure 3.5.

Left Endpoint	Critical Number	Critical Number	Right Endpoint
$f(-1) = 7$	$f(0) = 0$	$f(1) = -1$ Minimum	$f(2) = 16$ Maximum

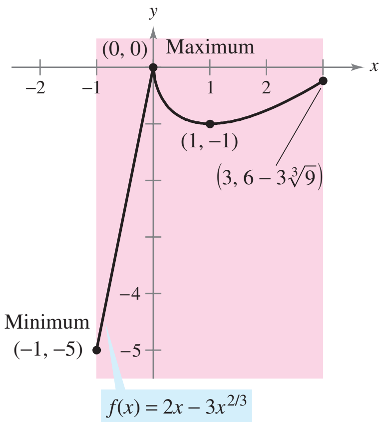


On the closed interval $[-1, 2]$, f has a minimum at $(1, -1)$ and a maximum at $(2, 16)$.
Figure 3.5

Editable Graph

Try It Exploration A Exploration B Video

In Figure 3.5, note that the critical number $x = 0$ does not yield a relative minimum or a relative maximum. This tells you that the converse of Theorem 3.2 is not true. In other words, *the critical numbers of a function need not produce relative extrema.*



On the closed interval $[-1, 3]$, f has a minimum at $(-1, -5)$ and a maximum at $(0, 0)$.
Figure 3.6

Editable Graph

EXAMPLE 3 Finding Extrema on a Closed Interval

Find the extrema of $f(x) = 2x - 3x^{2/3}$ on the interval $[-1, 3]$.

Solution Begin by differentiating the function.

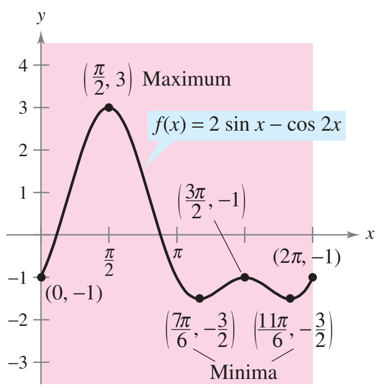
$$f(x) = 2x - 3x^{2/3} \qquad \text{Write original function.}$$

$$f'(x) = 2 - \frac{2}{x^{1/3}} = 2\left(\frac{x^{1/3} - 1}{x^{1/3}}\right) \qquad \text{Differentiate.}$$

From this derivative, you can see that the function has two critical numbers in the interval $[-1, 3]$. The number 1 is a critical number because $f'(1) = 0$, and the number 0 is a critical number because $f'(0)$ does not exist. By evaluating f at these two numbers and at the endpoints of the interval, you can conclude that the minimum is $f(-1) = -5$ and the maximum is $f(0) = 0$, as shown in the table. The graph of f is shown in Figure 3.6.

Left Endpoint	Critical Number	Critical Number	Right Endpoint
$f(-1) = -5$ Minimum	$f(0) = 0$ Maximum	$f(1) = -1$	$f(3) = 6 - 3\sqrt[3]{9} \approx -0.24$

Try It　Exploration A



On the closed interval $[0, 2\pi]$, f has two minima at $(7\pi/6, -3/2)$ and $(11\pi/6, -3/2)$ and a maximum at $(\pi/2, 3)$.
Figure 3.7

Editable Graph

EXAMPLE 4 Finding Extrema on a Closed Interval

Find the extrema of $f(x) = 2\sin x - \cos 2x$ on the interval $[0, 2\pi]$.

Solution This function is differentiable for all real x, so you can find all critical numbers by differentiating the function and setting $f'(x)$ equal to zero, as shown.

$$f(x) = 2\sin x - \cos 2x \qquad \text{Write original function.}$$

$$f'(x) = 2\cos x + 2\sin 2x = 0 \qquad \text{Set } f'(x) \text{ equal to 0.}$$

$$2\cos x + 4\cos x \sin x = 0 \qquad \sin 2x = 2\cos x \sin x$$

$$2(\cos x)(1 + 2\sin x) = 0 \qquad \text{Factor.}$$

In the interval $[0, 2\pi]$, the factor $\cos x$ is zero when $x = \pi/2$ and when $x = 3\pi/2$. The factor $(1 + 2\sin x)$ is zero when $x = 7\pi/6$ and when $x = 11\pi/6$. By evaluating f at these four critical numbers and at the endpoints of the interval, you can conclude that the maximum is $f(\pi/2) = 3$ and the minimum occurs at *two* points, $f(7\pi/6) = -3/2$ and $f(11\pi/6) = -3/2$, as shown in the table. The graph is shown in Figure 3.7.

Left Endpoint	Critical Number	Critical Number	Critical Number	Critical Number	Right Endpoint
$f(0) = -1$	$f\left(\frac{\pi}{2}\right) = 3$ Maximum	$f\left(\frac{7\pi}{6}\right) = -\frac{3}{2}$ Minimum	$f\left(\frac{3\pi}{2}\right) = -1$	$f\left(\frac{11\pi}{6}\right) = -\frac{3}{2}$ Minimum	$f(2\pi) = -1$

Try It　Open Exploration

Section 3.2 Rolle's Theorem and the Mean Value Theorem

- **Understand and use Rolle's Theorem.**
- **Understand and use the Mean Value Theorem.**

Rolle's Theorem

The Extreme Value Theorem (Section 3.1) states that a continuous function on a closed interval $[a, b]$ must have both a minimum and a maximum on the interval. Both of these values, however, can occur at the endpoints. **Rolle's Theorem,** named after the French mathematician Michel Rolle (1652–1719), gives conditions that guarantee the existence of an extreme value in the *interior* of a closed interval.

ROLLE'S THEOREM

French mathematician Michel Rolle first published the theorem that bears his name in 1691. Before this time, however, Rolle was one of the most vocal critics of calculus, stating that it gave erroneous results and was based on unsound reasoning. Later in life, Rolle came to see the usefulness of calculus.

EXPLORATION

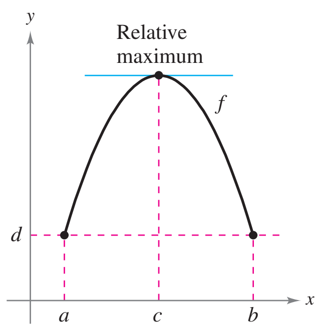
Extreme Values in a Closed Interval Sketch a rectangular coordinate plane on a piece of paper. Label the points $(1, 3)$ and $(5, 3)$. Using a pencil or pen, draw the graph of a differentiable function f that starts at $(1, 3)$ and ends at $(5, 3)$. Is there at least one point on the graph for which the derivative is zero? Would it be possible to draw the graph so that there *isn't* a point for which the derivative is zero? Explain your reasoning.

THEOREM 3.3 Rolle's Theorem

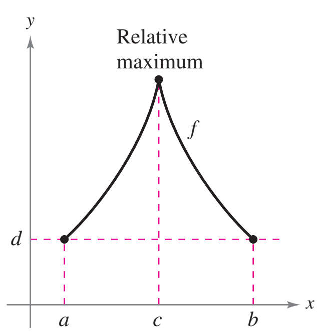
Let f be continuous on the closed interval $[a, b]$ and differentiable on the open interval (a, b). If

$$f(a) = f(b)$$

then there is at least one number c in (a, b) such that $f'(c) = 0$.



(a) f is continuous on $[a, b]$ and differentiable on (a, b).



(b) f is continuous on $[a, b]$.
Figure 3.8

Proof Let $f(a) = d = f(b)$.

Case 1: If $f(x) = d$ for all x in $[a, b]$, f is constant on the interval and, by Theorem 2.2, $f'(x) = 0$ for all x in (a, b).

Case 2: Suppose $f(x) > d$ for some x in (a, b). By the Extreme Value Theorem, you know that f has a maximum at some c in the interval. Moreover, because $f(c) > d$, this maximum does not occur at either endpoint. So, f has a maximum in the *open* interval (a, b). This implies that $f(c)$ is a *relative* maximum and, by Theorem 3.2, c is a critical number of f. Finally, because f is differentiable at c, you can conclude that $f'(c) = 0$.

Case 3: If $f(x) < d$ for some x in (a, b), you can use an argument similar to that in Case 2, but involving the minimum instead of the maximum.

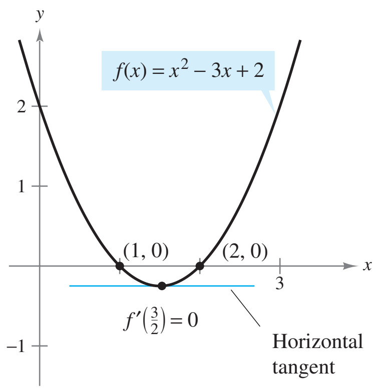
From Rolle's Theorem, you can see that if a function f is continuous on $[a, b]$ and differentiable on (a, b), and if $f(a) = f(b)$, there must be at least one x-value between a and b at which the graph of f has a horizontal tangent, as shown in Figure 3.8(a). If the differentiability requirement is dropped from Rolle's Theorem, f will still have a critical number in (a, b), but it may not yield a horizontal tangent. Such a case is shown in Figure 3.8(b).

EXAMPLE 1 Illustrating Rolle's Theorem

Find the two x-intercepts of

$$f(x) = x^2 - 3x + 2$$

and show that $f'(x) = 0$ at some point between the two x-intercepts.



The x-value for which $f'(x) = 0$ is between the two x-intercepts.
Figure 3.9

Editable Graph

Solution Note that f is differentiable on the entire real line. Setting $f(x)$ equal to 0 produces

$$x^2 - 3x + 2 = 0 \qquad \text{Set } f(x) \text{ equal to 0.}$$
$$(x - 1)(x - 2) = 0. \qquad \text{Factor.}$$

So, $f(1) = f(2) = 0$, and from Rolle's Theorem you know that there *exists* at least one c in the interval $(1, 2)$ such that $f'(c) = 0$. To *find* such a c, you can solve the equation

$$f'(x) = 2x - 3 = 0 \qquad \text{Set } f'(x) \text{ equal to 0.}$$

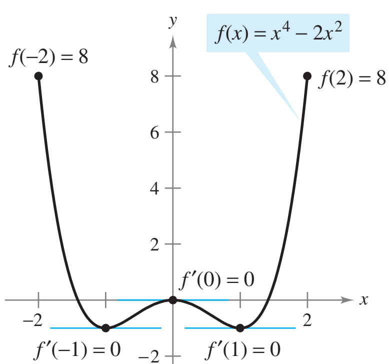
and determine that $f'(x) = 0$ when $x = \frac{3}{2}$. Note that the x-value lies in the open interval $(1, 2)$, as shown in Figure 3.9.

Try It Exploration A

Rolle's Theorem states that if f satisfies the conditions of the theorem, there must be *at least* one point between a and b at which the derivative is 0. There may of course be more than one such point, as shown in the next example.

EXAMPLE 2 Illustrating Rolle's Theorem

Let $f(x) = x^4 - 2x^2$. Find all values of c in the interval $(-2, 2)$ such that $f'(c) = 0$.



$f'(x) = 0$ for more than one x-value in the interval $(-2, 2)$.
Figure 3.10

Editable Graph

Solution To begin, note that the function satisfies the conditions of Rolle's Theorem. That is, f is continuous on the interval $[-2, 2]$ and differentiable on the interval $(-2, 2)$. Moreover, because $f(-2) = f(2) = 8$, you can conclude that there exists at least one c in $(-2, 2)$ such that $f'(c) = 0$. Setting the derivative equal to 0 produces

$$f'(x) = 4x^3 - 4x = 0 \qquad \text{Set } f'(x) \text{ equal to 0.}$$
$$4x(x - 1)(x + 1) = 0 \qquad \text{Factor.}$$
$$x = 0, 1, -1. \qquad x\text{-values for which } f'(x) = 0$$

So, in the interval $(-2, 2)$, the derivative is zero at three different values of x, as shown in Figure 3.10.

Try It Exploration A

TECHNOLOGY PITFALL A graphing utility can be used to indicate whether the points on the graphs in Examples 1 and 2 are relative minima or relative maxima of the functions. When using a graphing utility, however, you should keep in mind that it can give misleading pictures of graphs. For example, use a graphing utility to graph

$$f(x) = 1 - (x - 1)^2 - \frac{1}{1000(x - 1)^{1/7} + 1}.$$

With most viewing windows, it appears that the function has a maximum of 1 when $x = 1$ (see Figure 3.11). By evaluating the function at $x = 1$, however, you can see that $f(1) = 0$. To determine the behavior of this function near $x = 1$, you need to examine the graph analytically to get the complete picture.

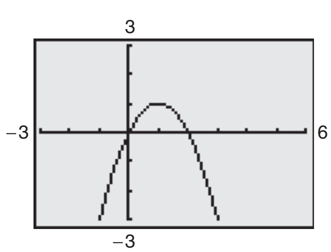


Figure 3.11

The Mean Value Theorem

Rolle's Theorem can be used to prove another theorem—the **Mean Value Theorem.**

THEOREM 3.4 The Mean Value Theorem

If f is continuous on the closed interval $[a, b]$ and differentiable on the open interval (a, b), then there exists a number c in (a, b) such that

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

Video

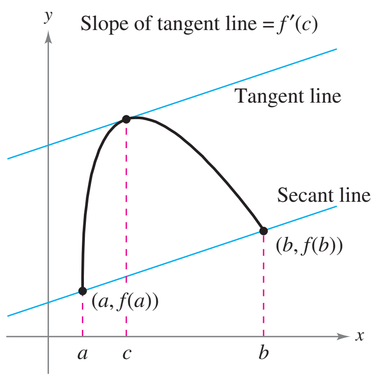


Figure 3.12

Proof Refer to Figure 3.12. The equation of the secant line containing the points $(a, f(a))$ and $(b, f(b))$ is

$$y = \left[\frac{f(b) - f(a)}{b - a}\right](x - a) + f(a).$$

Let $g(x)$ be the difference between $f(x)$ and y. Then

$$\begin{aligned} g(x) &= f(x) - y \\ &= f(x) - \left[\frac{f(b) - f(a)}{b - a}\right](x - a) - f(a). \end{aligned}$$

By evaluating g at a and b, you can see that $g(a) = 0 = g(b)$. Because f is continuous on $[a, b]$ it follows that g is also continuous on $[a, b]$. Furthermore, because f is differentiable, g is also differentiable, and you can apply Rolle's Theorem to the function g. So, there exists a number c in (a, b) such that $g'(c) = 0$, which implies that

$$\begin{aligned} 0 &= g'(c) \\ &= f'(c) - \frac{f(b) - f(a)}{b - a}. \end{aligned}$$

So, there exists a number c in (a, b) such that

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

NOTE The "mean" in the Mean Value Theorem refers to the mean (or average) rate of change of f in the interval $[a, b]$.

Joseph-Louis Lagrange (1736–1813)

The Mean Value Theorem was first proved by the famous mathematician Joseph-Louis Lagrange. Born in Italy, Lagrange held a position in the court of Frederick the Great in Berlin for 20 years. Afterward, he moved to France, where he met emperor Napoleon Bonaparte, who is quoted as saying, "Lagrange is the lofty pyramid of the mathematical sciences."

MathBio

Although the Mean Value Theorem can be used directly in problem solving, it is used more often to prove other theorems. In fact, some people consider this to be the most important theorem in calculus—it is closely related to the Fundamental Theorem of Calculus discussed in Chapter 4. For now, you can get an idea of the versatility of this theorem by looking at the results stated in Exercises 77–85 in this section.

The Mean Value Theorem has implications for both basic interpretations of the derivative. Geometrically, the theorem guarantees the existence of a tangent line that is parallel to the secant line through the points $(a, f(a))$ and $(b, f(b))$, as shown in Figure 3.12. Example 3 illustrates this geometric interpretation of the Mean Value Theorem. In terms of rates of change, the Mean Value Theorem implies that there must be a point in the open interval (a, b) at which the instantaneous rate of change is equal to the average rate of change over the interval $[a, b]$. This is illustrated in Example 4.

EXAMPLE 3 Finding a Tangent Line

Given $f(x) = 5 - (4/x)$, find all values of c in the open interval $(1, 4)$ such that

$$f'(c) = \frac{f(4) - f(1)}{4 - 1}.$$

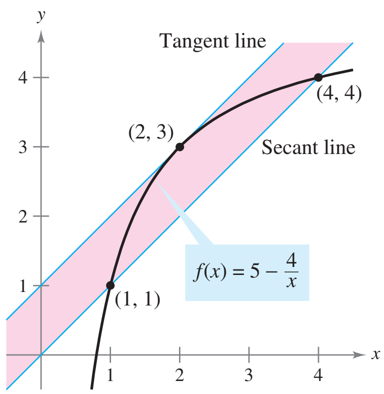
Solution The slope of the secant line through $(1, f(1))$ and $(4, f(4))$ is

$$\frac{f(4) - f(1)}{4 - 1} = \frac{4 - 1}{4 - 1} = 1.$$

Because f satisfies the conditions of the Mean Value Theorem, there exists at least one number c in $(1, 4)$ such that $f'(c) = 1$. Solving the equation $f'(x) = 1$ yields

$$f'(x) = \frac{4}{x^2} = 1$$

which implies that $x = \pm 2$. So, in the interval $(1, 4)$, you can conclude that $c = 2$, as shown in Figure 3.13.



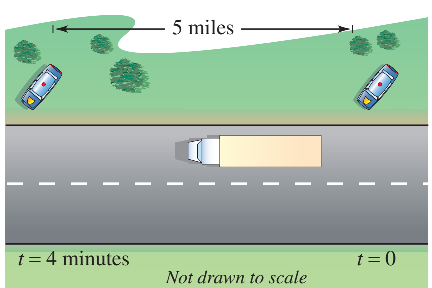
The tangent line at $(2, 3)$ is parallel to the secant line through $(1, 1)$ and $(4, 4)$.
Figure 3.13

Editable Graph

Try It | Exploration A | Open Exploration | Video

EXAMPLE 4 Finding an Instantaneous Rate of Change

Two stationary patrol cars equipped with radar are 5 miles apart on a highway, as shown in Figure 3.14. As a truck passes the first patrol car, its speed is clocked at 55 miles per hour. Four minutes later, when the truck passes the second patrol car, its speed is clocked at 50 miles per hour. Prove that the truck must have exceeded the speed limit (of 55 miles per hour) at some time during the 4 minutes.



At some time t, the instantaneous velocity is equal to the average velocity over 4 minutes.
Figure 3.14

Animation

Solution Let $t = 0$ be the time (in hours) when the truck passes the first patrol car. The time when the truck passes the second patrol car is

$$t = \frac{4}{60} = \frac{1}{15} \text{ hour.}$$

By letting $s(t)$ represent the distance (in miles) traveled by the truck, you have $s(0) = 0$ and $s\left(\frac{1}{15}\right) = 5$. So, the average velocity of the truck over the five-mile stretch of highway is

$$\begin{aligned} \text{Average velocity} &= \frac{s(1/15) - s(0)}{(1/15) - 0} \\ &= \frac{5}{1/15} = 75 \text{ miles per hour.} \end{aligned}$$

Assuming that the position function is differentiable, you can apply the Mean Value Theorem to conclude that the truck must have been traveling at a rate of 75 miles per hour sometime during the 4 minutes.

Try It | Exploration A

A useful alternative form of the Mean Value Theorem is as follows: If f is continuous on $[a, b]$ and differentiable on (a, b), then there exists a number c in (a, b) such that

$$f(b) = f(a) + (b - a)f'(c).$$ Alternative form of Mean Value Theorem

NOTE When doing the exercises for this section, keep in mind that polynomial functions, rational functions, and trigonometric functions are differentiable at all points in their domains.

Section 3.3 Increasing and Decreasing Functions and the First Derivative Test

- **Determine intervals on which a function is increasing or decreasing.**
- **Apply the First Derivative Test to find relative extrema of a function.**

Increasing and Decreasing Functions

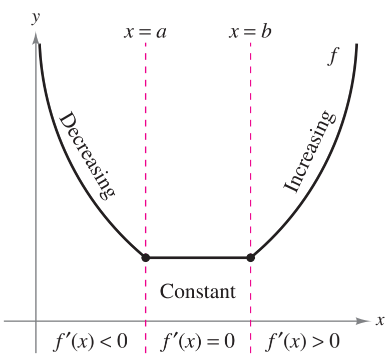
In this section you will learn how derivatives can be used to *classify* relative extrema as either relative minima or relative maxima. First, it is important to define increasing and decreasing functions.

Definitions of Increasing and Decreasing Functions

A function f is **increasing** on an interval if for any two numbers x_1 and x_2 in the interval, $x_1 < x_2$ implies $f(x_1) < f(x_2)$.

A function f is **decreasing** on an interval if for any two numbers x_1 and x_2 in the interval, $x_1 < x_2$ implies $f(x_1) > f(x_2)$.

Video



The derivative is related to the slope of a function.
Figure 3.15

A function is increasing if, *as x moves to the right*, its graph moves up, and is decreasing if its graph moves down. For example, the function in Figure 3.15 is decreasing on the interval $(-\infty, a)$, is constant on the interval (a, b), and is increasing on the interval (b, ∞). As shown in Theorem 3.5 below, a positive derivative implies that the function is increasing; a negative derivative implies that the function is decreasing; and a zero derivative on an entire interval implies that the function is constant on that interval.

THEOREM 3.5 Test for Increasing and Decreasing Functions

Let f be a function that is continuous on the closed interval $[a, b]$ and differentiable on the open interval (a, b).

1. If $f'(x) > 0$ for all x in (a, b), then f is increasing on $[a, b]$.
2. If $f'(x) < 0$ for all x in (a, b), then f is decreasing on $[a, b]$.
3. If $f'(x) = 0$ for all x in (a, b), then f is constant on $[a, b]$.

Proof To prove the first case, assume that $f'(x) > 0$ for all x in the interval (a, b) and let $x_1 < x_2$ be any two points in the interval. By the Mean Value Theorem, you know that there exists a number c such that $x_1 < c < x_2$, and

$$f'(c) = \frac{f(x_2) - f(x_1)}{x_2 - x_1}.$$

Because $f'(c) > 0$ and $x_2 - x_1 > 0$, you know that

$$f(x_2) - f(x_1) > 0$$

which implies that $f(x_1) < f(x_2)$. So, f is increasing on the interval. The second case has a similar proof (see Exercise 101), and the third case was given as Exercise 78 in Section 3.2.

NOTE The conclusions in the first two cases of Theorem 3.5 are valid even if $f'(x) = 0$ at a finite number of x-values in (a, b).

EXAMPLE 1 Intervals on Which f Is Increasing or Decreasing

Find the open intervals on which $f(x) = x^3 - \frac{3}{2}x^2$ is increasing or decreasing.

Solution Note that f is differentiable on the entire real number line. To determine the critical numbers of f, set $f'(x)$ equal to zero.

$$f(x) = x^3 - \frac{3}{2}x^2 \qquad \text{Write original function.}$$

$$f'(x) = 3x^2 - 3x = 0 \qquad \text{Differentiate and set } f'(x) \text{ equal to 0.}$$

$$3(x)(x-1) = 0 \qquad \text{Factor.}$$

$$x = 0, 1 \qquad \text{Critical numbers}$$

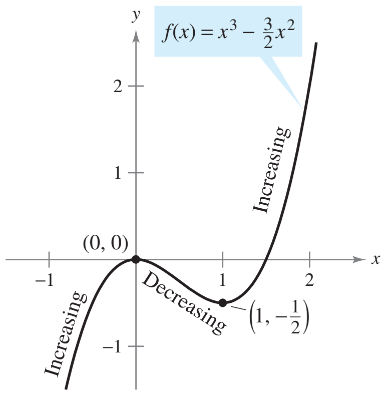


Figure 3.16

Editable Graph

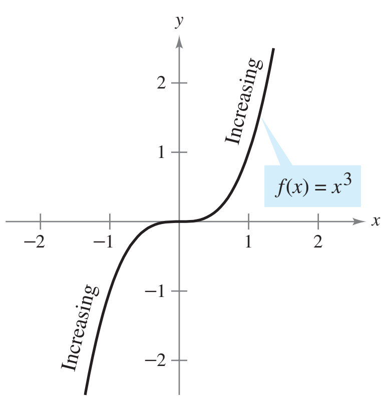
Because there are no points for which f' does not exist, you can conclude that $x = 0$ and $x = 1$ are the only critical numbers. The table summarizes the testing of the three intervals determined by these two critical numbers.

Interval	$-\infty < x < 0$	$0 < x < 1$	$1 < x < \infty$
Test Value	$x = -1$	$x = \frac{1}{2}$	$x = 2$
Sign of $f'(x)$	$f'(-1) = 6 > 0$	$f'\left(\frac{1}{2}\right) = -\frac{3}{4} < 0$	$f'(2) = 6 > 0$
Conclusion	Increasing	Decreasing	Increasing

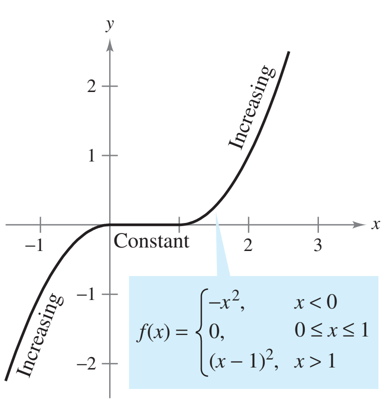
So, f is increasing on the intervals $(-\infty, 0)$ and $(1, \infty)$ and decreasing on the interval $(0, 1)$, as shown in Figure 3.16.

Try It Exploration A Exploration B Video

Example 1 gives you one example of how to find intervals on which a function is increasing or decreasing. The guidelines below summarize the steps followed in the example.



(a) Strictly monotonic function



(b) Not strictly monotonic

Figure 3.17

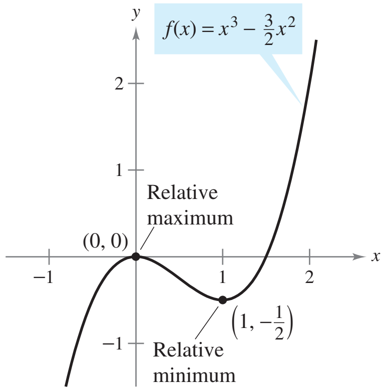
Guidelines for Finding Intervals on Which a Function Is Increasing or Decreasing

Let f be continuous on the interval (a, b). To find the open intervals on which f is increasing or decreasing, use the following steps.

1. Locate the critical numbers of f in (a, b), and use these numbers to determine test intervals.
2. Determine the sign of $f'(x)$ at one test value in each of the intervals.
3. Use Theorem 3.5 to determine whether f is increasing or decreasing on each interval.

These guidelines are also valid if the interval (a, b) is replaced by an interval of the form $(-\infty, b)$, (a, ∞), or $(-\infty, \infty)$.

A function is **strictly monotonic** on an interval if it is either increasing on the entire interval or decreasing on the entire interval. For instance, the function $f(x) = x^3$ is strictly monotonic on the entire real line because it is increasing on the entire real line, as shown in Figure 3.17(a). The function shown in Figure 3.17(b) is not strictly monotonic on the entire real line because it is constant on the interval $[0, 1]$.



Relative extrema of f

Figure 3.18

The First Derivative Test

After you have determined the intervals on which a function is increasing or decreasing, it is not difficult to locate the relative extrema of the function. For instance, in Figure 3.18 (from Example 1), the function

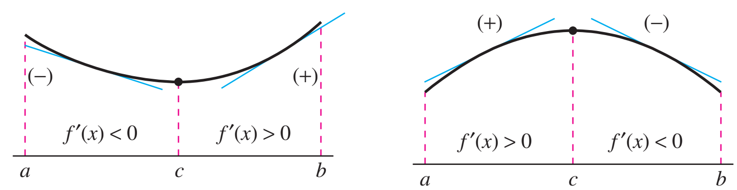
$$f(x) = x^3 - \frac{3}{2}x^2$$

has a relative maximum at the point $(0, 0)$ because f is increasing immediately to the left of $x = 0$ and decreasing immediately to the right of $x = 0$. Similarly, f has a relative minimum at the point $\left(1, -\frac{1}{2}\right)$ because f is decreasing immediately to the left of $x = 1$ and increasing immediately to the right of $x = 1$. The following theorem, called the First Derivative Test, makes this more explicit.

THEOREM 3.6 The First Derivative Test

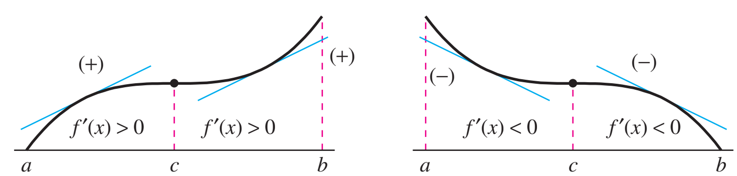
Let c be a critical number of a function f that is continuous on an open interval I containing c. If f is differentiable on the interval, except possibly at c, then $f(c)$ can be classified as follows.

1. If $f'(x)$ changes from negative to positive at c, then f has a *relative minimum* at $(c, f(c))$.
2. If $f'(x)$ changes from positive to negative at c, then f has a *relative maximum* at $(c, f(c))$.
3. If $f'(x)$ is positive on both sides of c or negative on both sides of c, then $f(c)$ is neither a relative minimum nor a relative maximum.



Relative minimum

Relative maximum



Neither relative minimum nor relative maximum

Proof Assume that $f'(x)$ changes from negative to positive at c. Then there exist a and b in I such that

$$f'(x) < 0 \text{ for all } x \text{ in } (a, c)$$

and

$$f'(x) > 0 \text{ for all } x \text{ in } (c, b).$$

By Theorem 3.5, f is decreasing on (a, c) and increasing on (c, b). So, $f(c)$ is a minimum of f on the open interval (a, b) and, consequently, a relative minimum of f. This proves the first case of the theorem. The second case can be proved in a similar way (see Exercise 102).

Video

EXAMPLE 2 Applying the First Derivative Test

Find the relative extrema of the function $f(x) = \frac{1}{2}x - \sin x$ in the interval $(0, 2\pi)$.

Solution Note that f is continuous on the interval $(0, 2\pi)$. To determine the critical numbers of f in this interval, set $f'(x)$ equal to 0.

$$f'(x) = \frac{1}{2} - \cos x = 0 \qquad \text{Set } f'(x) \text{ equal to 0.}$$

$$\cos x = \frac{1}{2}$$

$$x = \frac{\pi}{3}, \frac{5\pi}{3} \qquad \text{Critical numbers}$$

Because there are no points for which f' does not exist, you can conclude that $x = \pi/3$ and $x = 5\pi/3$ are the only critical numbers. The table summarizes the testing of the three intervals determined by these two critical numbers.

Interval	$0 < x < \frac{\pi}{3}$	$\frac{\pi}{3} < x < \frac{5\pi}{3}$	$\frac{5\pi}{3} < x < 2\pi$
Test Value	$x = \frac{\pi}{4}$	$x = \pi$	$x = \frac{7\pi}{4}$
Sign of $f'(x)$	$f'\left(\frac{\pi}{4}\right) < 0$	$f'(\pi) > 0$	$f'\left(\frac{7\pi}{4}\right) < 0$
Conclusion	Decreasing	Increasing	Decreasing

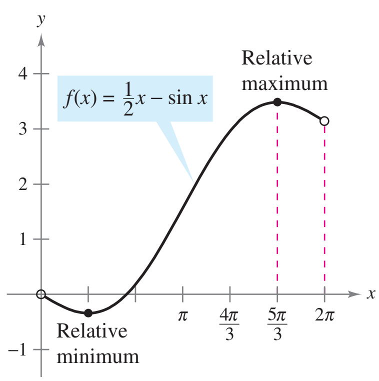
By applying the First Derivative Test, you can conclude that f has a relative minimum at the point where

$$x = \frac{\pi}{3} \qquad x\text{-value where relative minimum occurs}$$

and a relative maximum at the point where

$$x = \frac{5\pi}{3} \qquad x\text{-value where relative maximum occurs}$$

as shown in Figure 3.19.



A relative minimum occurs where f changes from decreasing to increasing, and a relative maximum occurs where f changes from increasing to decreasing.
Figure 3.19

Editable Graph

Try It | Exploration A | Exploration B

EXPLORATION

Comparing Graphical and Analytic Approaches From Section 3.2, you know that, *by itself*, a graphing utility can give misleading information about the relative extrema of a graph. *Used in conjunction with an analytic approach*, however, a graphing utility can provide a good way to reinforce your conclusions. Use a graphing utility to graph the function in Example 2. Then use the *zoom* and *trace* features to estimate the relative extrema. How close are your graphical approximations?

Note that in Examples 1 and 2 the given functions are differentiable on the entire real line. For such functions, the only critical numbers are those for which $f'(x) = 0$. Example 3 concerns a function that has two types of critical numbers—those for which $f'(x) = 0$ and those for which f is not differentiable.

EXAMPLE 3 Applying the First Derivative Test

Find the relative extrema of

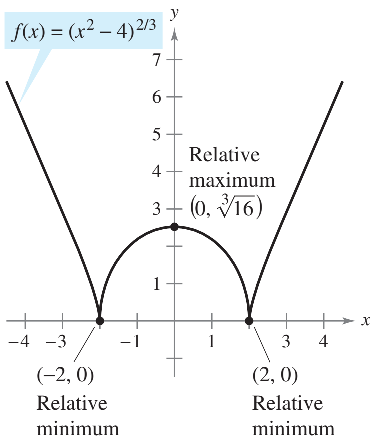
$$f(x) = (x^2 - 4)^{2/3}.$$

Solution Begin by noting that f is continuous on the entire real line. The derivative of f

$$f'(x) = \frac{2}{3}(x^2 - 4)^{-1/3}(2x) \qquad \text{General Power Rule}$$

$$= \frac{4x}{3(x^2 - 4)^{1/3}} \qquad \text{Simplify.}$$

is 0 when $x = 0$ and does not exist when $x = \pm 2$. So, the critical numbers are $x = -2$, $x = 0$, and $x = 2$. The table summarizes the testing of the four intervals determined by these three critical numbers.



You can apply the First Derivative Test to find relative extrema.
Figure 3.20

Editable Graph

Interval	$-\infty < x < -2$	$-2 < x < 0$	$0 < x < 2$	$2 < x < \infty$
Test Value	$x = -3$	$x = -1$	$x = 1$	$x = 3$
Sign of $f'(x)$	$f'(-3) < 0$	$f'(-1) > 0$	$f'(1) < 0$	$f'(3) > 0$
Conclusion	Decreasing	Increasing	Decreasing	Increasing

By applying the First Derivative Test, you can conclude that f has a relative minimum at the point $(-2, 0)$, a relative maximum at the point $(0, \sqrt[3]{16})$, and another relative minimum at the point $(2, 0)$, as shown in Figure 3.20.

Try It

Exploration A

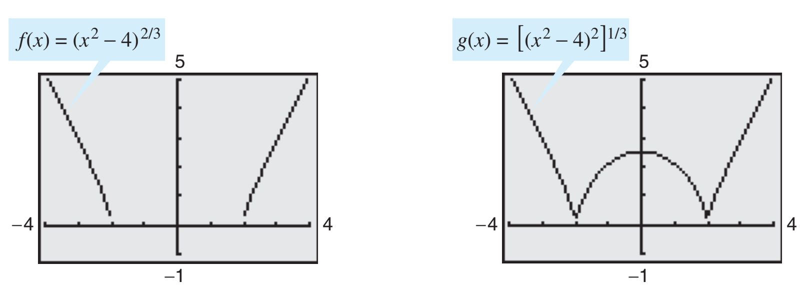
TECHNOLOGY PITFALL When using a graphing utility to graph a function involving radicals or rational exponents, be sure you understand the way the utility evaluates radical expressions. For instance, even though

$$f(x) = (x^2 - 4)^{2/3}$$

and

$$g(x) = [(x^2 - 4)^2]^{1/3}$$

are the same algebraically, some graphing utilities distinguish between these two functions. Which of the graphs shown in Figure 3.21 is incorrect? Why did the graphing utility produce an incorrect graph?



Which graph is incorrect?
Figure 3.21

When using the First Derivative Test, be sure to consider the domain of the function. For instance, in the next example, the function

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

is not defined when $x = 0$. This x-value must be used with the critical numbers to determine the test intervals.

EXAMPLE 4 Applying the First Derivative Test

Find the relative extrema of $f(x) = \dfrac{x^4 + 1}{x^2}$.

Solution

$$\begin{aligned} f(x) &= x^2 + x^{-2} && \text{Rewrite original function.} \\ f'(x) &= 2x - 2x^{-3} && \text{Differentiate.} \\ &= 2x - \frac{2}{x^3} && \text{Rewrite with positive exponent.} \\ &= \frac{2(x^4 - 1)}{x^3} && \text{Simplify.} \\ &= \frac{2(x^2 + 1)(x - 1)(x + 1)}{x^3} && \text{Factor.} \end{aligned}$$

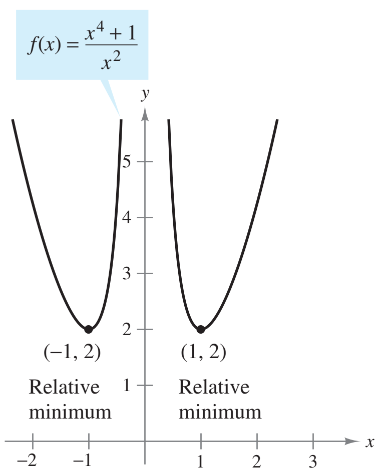
So, $f'(x)$ is zero at $x = \pm 1$. Moreover, because $x = 0$ is not in the domain of f, you should use this x-value along with the critical numbers to determine the test intervals.

$$\begin{aligned} x &= \pm 1 && \text{Critical numbers, } f'(\pm 1) = 0 \\ x &= 0 && 0 \text{ is not in the domain of } f. \end{aligned}$$

The table summarizes the testing of the four intervals determined by these three x-values.

Interval	$-\infty < x < -1$	$-1 < x < 0$	$0 < x < 1$	$1 < x < \infty$
Test Value	$x = -2$	$x = -\frac{1}{2}$	$x = \frac{1}{2}$	$x = 2$
Sign of $f'(x)$	$f'(-2) < 0$	$f'\left(-\frac{1}{2}\right) > 0$	$f'\left(\frac{1}{2}\right) < 0$	$f'(2) > 0$
Conclusion	Decreasing	Increasing	Decreasing	Increasing

By applying the First Derivative Test, you can conclude that f has one relative minimum at the point $(-1, 2)$ and another at the point $(1, 2)$, as shown in Figure 3.22.



x-values that are not in the domain of f, as well as critical numbers, determine test intervals for f'.
Figure 3.22

Editable Graph

Try It

Exploration A

Open Exploration

TECHNOLOGY The most difficult step in applying the First Derivative Test is finding the values for which the derivative is equal to 0. For instance, the values of x for which the derivative of

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

is equal to zero are $x = 0$ and $x = \pm\sqrt{\sqrt{2} - 1}$. If you have access to technology that can perform symbolic differentiation and solve equations, use it to apply the First Derivative Test to this function.

If a projectile is propelled from ground level and air resistance is neglected, the object will travel farthest with an initial angle of 45°. If, however, the projectile is propelled from a point above ground level, the angle that yields a maximum horizontal distance is not 45° (see Example 5).

EXAMPLE 5 The Path of a Projectile

Neglecting air resistance, the path of a projectile that is propelled at an angle θ is

$$y = \frac{g \sec^2 \theta}{2v_0^2}x^2 + (\tan \theta)x + h, \quad 0 \le \theta \le \frac{\pi}{2}$$

where y is the height, x is the horizontal distance, g is the acceleration due to gravity, v_0 is the initial velocity, and h is the initial height. (This equation is derived in Section 12.3.) Let $g = -32$ feet per second per second, $v_0 = 24$ feet per second, and $h = 9$ feet. What value of θ will produce a maximum horizontal distance?

Solution To find the distance the projectile travels, let $y = 0$, and use the Quadratic Formula to solve for x.

$$\frac{g \sec^2 \theta}{2v_0^2}x^2 + (\tan \theta)x + h = 0$$

$$\frac{-32 \sec^2 \theta}{2(24^2)}x^2 + (\tan \theta)x + 9 = 0$$

$$-\frac{\sec^2 \theta}{36}x^2 + (\tan \theta)x + 9 = 0$$

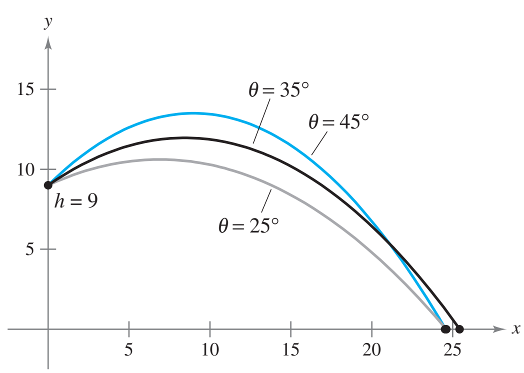
$$x = \frac{-\tan \theta \pm \sqrt{\tan^2 \theta + \sec^2 \theta}}{-\sec^2 \theta/18}$$

$$x = 18 \cos \theta\left(\sin \theta + \sqrt{\sin^2 \theta + 1}\right), \quad x \ge 0$$

At this point, you need to find the value of θ that produces a maximum value of x. Applying the First Derivative Test by hand would be very tedious. Using technology to solve the equation $dx/d\theta = 0$, however, eliminates most of the messy computations. The result is that the maximum value of x occurs when

$$\theta \approx 0.61548 \text{ radian, or } 35.3°.$$

This conclusion is reinforced by sketching the path of the projectile for different values of θ, as shown in Figure 3.23. Of the three paths shown, note that the distance traveled is greatest for $\theta = 35°$.



The path of a projectile with initial angle θ
Figure 3.23

Simulation

Try It

Section 3.4 Concavity and the Second Derivative Test

- **Determine intervals on which a function is concave upward or concave downward.**
- **Find any points of inflection of the graph of a function.**
- **Apply the Second Derivative Test to find relative extrema of a function.**

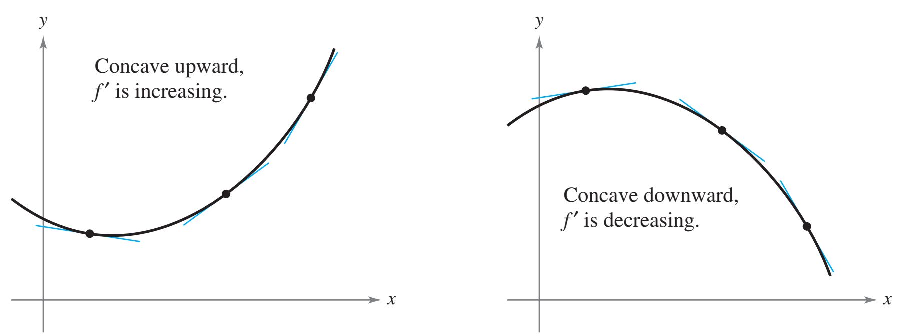
Concavity

You have already seen that locating the intervals in which a function f increases or decreases helps to describe its graph. In this section, you will see how locating the intervals in which f' increases or decreases can be used to determine where the graph of f is *curving upward* or *curving downward*.

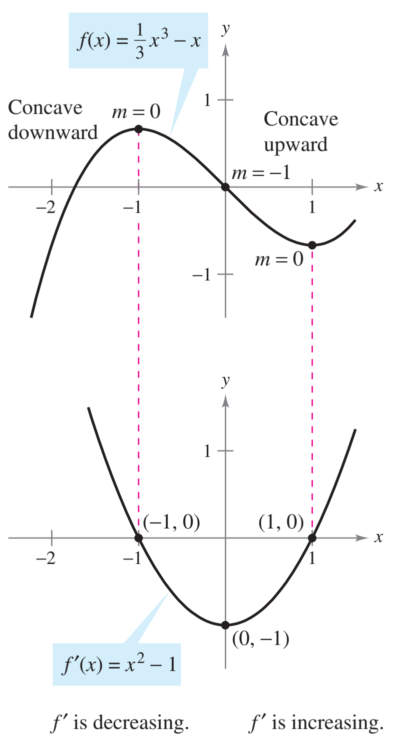
> **Definition of Concavity**
>
> Let f be differentiable on an open interval I. The graph of f is **concave upward** on I if f' is increasing on the interval and **concave downward** on I if f' is decreasing on the interval.

The following graphical interpretation of concavity is useful. (See Appendix A for a proof of these results.)

1. Let f be differentiable on an open interval I. If the graph of f is concave *upward* on I, then the graph of f lies *above* all of its tangent lines on I. [See Figure 3.24(a).]
2. Let f be differentiable on an open interval I. If the graph of f is concave *downward* on I, then the graph of f lies *below* all of its tangent lines on I. [See Figure 3.24(b).]



(a) The graph of f lies above its tangent lines. **(b)** The graph of f lies below its tangent lines.
Figure 3.24



The concavity of f is related to the slope of the derivative.
Figure 3.25

To find the open intervals on which the graph of a function f is concave upward or downward, you need to find the intervals on which f' is increasing or decreasing. For instance, the graph of

$$f(x) = \tfrac{1}{3}x^3 - x$$

is concave downward on the open interval $(-\infty, 0)$ because $f'(x) = x^2 - 1$ is decreasing there. (See Figure 3.25.) Similarly, the graph of f is concave upward on the interval $(0, \infty)$ because f' is increasing on $(0, \infty)$.

The following theorem shows how to use the *second* derivative of a function f to determine intervals on which the graph of f is concave upward or downward. A proof of this theorem follows directly from Theorem 3.5 and the definition of concavity.

THEOREM 3.7 Test for Concavity

Let f be a function whose second derivative exists on an open interval I.

1. If $f''(x) > 0$ for all x in I, then the graph of f is concave upward in I.
2. If $f''(x) < 0$ for all x in I, then the graph of f is concave downward in I.

NOTE A third case of Theorem 3.7 could be that if $f''(x) = 0$ for all x in I, then f is linear. Note, however, that concavity is not defined for a line. In other words, a straight line is neither concave upward nor concave downward.

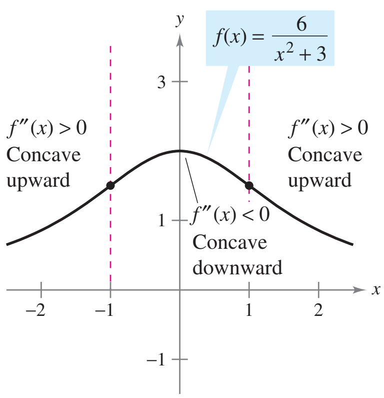
To apply Theorem 3.7, locate the x-values at which $f''(x) = 0$ or f'' does not exist. Second, use these x-values to determine test intervals. Finally, test the sign of $f''(x)$ in each of the test intervals.

EXAMPLE 1 Determining Concavity

Determine the open intervals on which the graph of

$$f(x) = \frac{6}{x^2 + 3}$$

is concave upward or downward.



From the sign of f'' you can determine the concavity of the graph of f.
Figure 3.26

Editable Graph

Solution Begin by observing that f is continuous on the entire real line. Next, find the second derivative of f.

$$\begin{aligned}
f(x) &= 6(x^2 + 3)^{-1} && \text{Rewrite original function.}\\
f'(x) &= (-6)(x^2 + 3)^{-2}(2x) && \text{Differentiate.}\\
&= \frac{-12x}{(x^2 + 3)^2} && \text{First derivative}\\
f''(x) &= \frac{(x^2 + 3)^2(-12) - (-12x)(2)(x^2 + 3)(2x)}{(x^2 + 3)^4} && \text{Differentiate.}\\
&= \frac{36(x^2 - 1)}{(x^2 + 3)^3} && \text{Second derivative}
\end{aligned}$$

Because $f''(x) = 0$ when $x = \pm 1$ and f'' is defined on the entire real line, you should test f'' in the intervals $(-\infty, -1)$, $(-1, 1)$, and $(1, \infty)$. The results are shown in the table and in Figure 3.26.

Interval	$-\infty < x < -1$	$-1 < x < 1$	$1 < x < \infty$
Test Value	$x = -2$	$x = 0$	$x = 2$
Sign of $f''(x)$	$f''(-2) > 0$	$f''(0) < 0$	$f''(2) > 0$
Conclusion	Concave upward	Concave downward	Concave upward

Try It Exploration A Video Video

The function given in Example 1 is continuous on the entire real line. If there are x-values at which the function is not continuous, these values should be used along with the points at which $f''(x) = 0$ or $f''(x)$ does not exist to form the test intervals.

EXAMPLE 2 Determining Concavity

Determine the open intervals on which the graph of $f(x) = \dfrac{x^2+1}{x^2-4}$ is concave upward or downward.

Solution Differentiating twice produces the following.

$$f(x) = \frac{x^2+1}{x^2-4} \qquad \text{Write original function.}$$

$$f'(x) = \frac{(x^2-4)(2x)-(x^2+1)(2x)}{(x^2-4)^2} \qquad \text{Differentiate.}$$

$$= \frac{-10x}{(x^2-4)^2} \qquad \text{First derivative}$$

$$f''(x) = \frac{(x^2-4)^2(-10)-(-10x)(2)(x^2-4)(2x)}{(x^2-4)^4} \qquad \text{Differentiate.}$$

$$= \frac{10(3x^2+4)}{(x^2-4)^3} \qquad \text{Second derivative}$$

There are no points at which $f''(x) = 0$, but at $x = \pm 2$ the function f is not continuous, so test for concavity in the intervals $(-\infty, -2)$, $(-2, 2)$, and $(2, \infty)$, as shown in the table. The graph of f is shown in Figure 3.27.

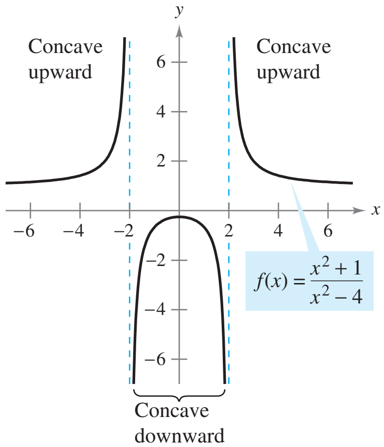


Figure 3.27

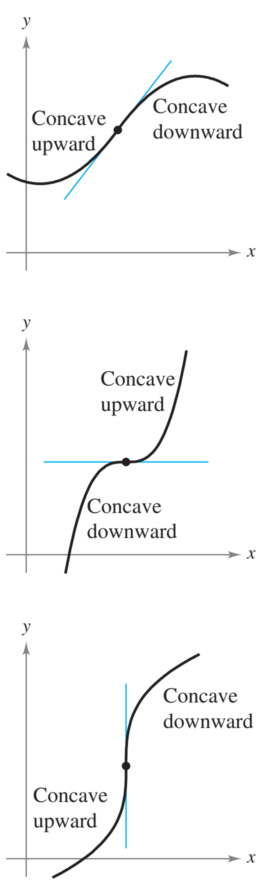
Editable Graph

Interval	$-\infty < x < -2$	$-2 < x < 2$	$2 < x < \infty$
Test Value	$x = -3$	$x = 0$	$x = 3$
Sign of $f''(x)$	$f''(-3) > 0$	$f''(0) < 0$	$f''(3) > 0$
Conclusion	Concave upward	Concave downward	Concave upward

Try It Exploration A Exploration B

Points of Inflection

The graph in Figure 3.26 has two points at which the concavity changes. If the tangent line to the graph exists at such a point, that point is a **point of inflection.** Three types of points of inflection are shown in Figure 3.28.



The concavity of f changes at a point of inflection. Note that a graph crosses its tangent line at a point of inflection.
Figure 3.28

Definition of Point of Inflection

Let f be a function that is continuous on an open interval and let c be a point in the interval. If the graph of f has a tangent line at this point $(c, f(c))$, then this point is a **point of inflection** of the graph of f if the concavity of f changes from upward to downward (or downward to upward) at the point.

NOTE The definition of *point of inflection* given in this book requires that the tangent line exists at the point of inflection. Some books do not require this. For instance, we do not consider the function

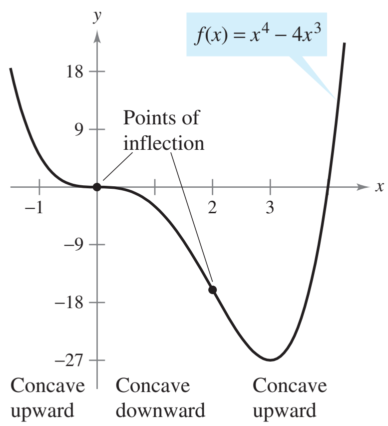
$$f(x) = \begin{cases} x^3, & x < 0 \\ x^2 + 2x, & x \ge 0 \end{cases}$$

to have a point of inflection at the origin, even though the concavity of the graph changes from concave downward to concave upward.

To locate *possible* points of inflection, you can determine the values of x for which $f''(x) = 0$ or $f''(x)$ does not exist. This is similar to the procedure for locating relative extrema of f.

THEOREM 3.8 Points of Inflection

If $(c, f(c))$ is a point of inflection of the graph of f, then either $f''(c) = 0$ or f'' does not exist at $x = c$.



Points of inflection can occur where $f''(x) = 0$ or f'' does not exist.
Figure 3.29

Editable Graph

EXAMPLE 3 Finding Points of Inflection

Determine the points of inflection and discuss the concavity of the graph of $f(x) = x^4 - 4x^3$.

Solution Differentiating twice produces the following.

$$f(x) = x^4 - 4x^3 \qquad \text{Write original function.}$$
$$f'(x) = 4x^3 - 12x^2 \qquad \text{Find first derivative.}$$
$$f''(x) = 12x^2 - 24x = 12x(x - 2) \qquad \text{Find second derivative.}$$

Setting $f''(x) = 0$, you can determine that the possible points of inflection occur at $x = 0$ and $x = 2$. By testing the intervals determined by these x-values, you can conclude that they both yield points of inflection. A summary of this testing is shown in the table, and the graph of f is shown in Figure 3.29.

Interval	$-\infty < x < 0$	$0 < x < 2$	$2 < x < \infty$
Test Value	$x = -1$	$x = 1$	$x = 3$
Sign of $f''(x)$	$f''(-1) > 0$	$f''(1) < 0$	$f''(3) > 0$
Conclusion	Concave upward	Concave downward	Concave upward

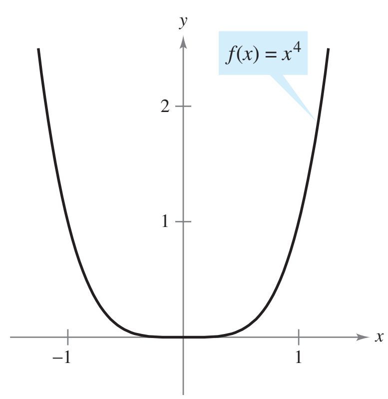
Try It

Exploration A

Exploration B

Exploration C

Exploration D



$f''(0) = 0$, but $(0, 0)$ is not a point of inflection.
Figure 3.30

The converse of Theorem 3.8 is not generally true. That is, it is possible for the second derivative to be 0 at a point that is *not* a point of inflection. For instance, the graph of $f(x) = x^4$ is shown in Figure 3.30. The second derivative is 0 when $x = 0$, but the point $(0, 0)$ is not a point of inflection because the graph of f is concave upward in both intervals $-\infty < x < 0$ and $0 < x < \infty$.

EXPLORATION

Consider a general cubic function of the form

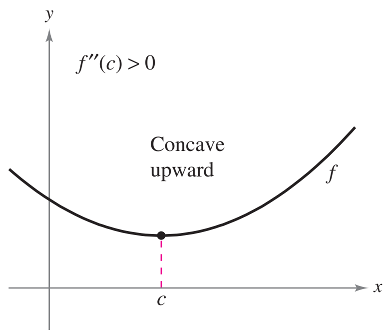
$$f(x) = ax^3 + bx^2 + cx + d.$$

You know that the value of d has a bearing on the location of the graph but has no bearing on the value of the first derivative at given values of x. Graphically, this is true because changes in the value of d shift the graph up or down but do not change its basic shape. Use a graphing utility to graph several cubics with different values of c. Then give a graphical explanation of why changes in c do not affect the values of the second derivative.

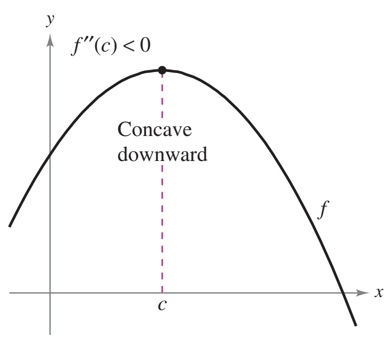
The Second Derivative Test

In addition to testing for concavity, the second derivative can be used to perform a simple test for relative maxima and minima. The test is based on the fact that if the graph of a function f is concave upward on an open interval containing c, and $f'(c) = 0$, $f(c)$ must be a relative minimum of f. Similarly, if the graph of a function f is concave downward on an open interval containing c, and $f'(c) = 0$, $f(c)$ must be a relative maximum of f (see Figure 3.31).

Video



If $f'(c) = 0$ and $f''(c) > 0$, $f(c)$ is a relative minimum.



If $f'(c) = 0$ and $f''(c) < 0$, $f(c)$ is a relative maximum.

Figure 3.31

THEOREM 3.9 Second Derivative Test

Let f be a function such that $f'(c) = 0$ and the second derivative of f exists on an open interval containing c.

1. If $f''(c) > 0$, then f has a relative minimum at $(c, f(c))$.
2. If $f''(c) < 0$, then f has a relative maximum at $(c, f(c))$.

If $f''(c) = 0$, the test fails. That is, f may have a relative maximum at c, a relative minimum at $(c, f(c))$, or neither. In such cases, you can use the First Derivative Test.

Proof If $f'(c) = 0$ and $f''(c) > 0$, there exists an open interval I containing c for which

$$\frac{f'(x) - f'(c)}{x - c} = \frac{f'(x)}{x - c} > 0$$

for all $x \neq c$ in I. If $x < c$, then $x - c < 0$ and $f'(x) < 0$. Also, if $x > c$, then $x - c > 0$ and $f'(x) > 0$. So, $f'(x)$ changes from negative to positive at c, and the First Derivative Test implies that $f(c)$ is a relative minimum. A proof of the second case is left to you.

EXAMPLE 4 Using the Second Derivative Test

Find the relative extrema for $f(x) = -3x^5 + 5x^3$.

Solution Begin by finding the critical numbers of f.

$$f'(x) = -15x^4 + 15x^2 = 15x^2(1 - x^2) = 0 \qquad \text{Set } f'(x) \text{ equal to 0.}$$

$$x = -1, 0, 1 \qquad \text{Critical numbers}$$

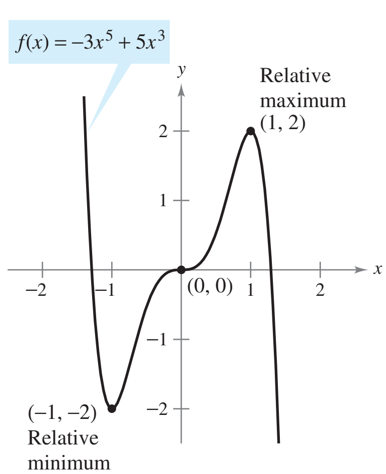
Using

$$f''(x) = -60x^3 + 30x = 30(-2x^3 + x)$$

you can apply the Second Derivative Test as shown below.

Point	$(-1, -2)$	$(1, 2)$	$(0, 0)$
Sign of $f''(x)$	$f''(-1) > 0$	$f''(1) < 0$	$f''(0) = 0$
Conclusion	Relative minimum	Relative maximum	Test fails

Because the Second Derivative Test fails at $(0, 0)$, you can use the First Derivative Test and observe that f increases to the left and right of $x = 0$. So, $(0, 0)$ is neither a relative minimum nor a relative maximum (even though the graph has a horizontal tangent line at this point). The graph of f is shown in Figure 3.32.



$(0, 0)$ is neither a relative minimum nor a relative maximum.

Figure 3.32

Editable Graph

Try It

Exploration A

Open Exploration

Section 3.5 Limits at Infinity

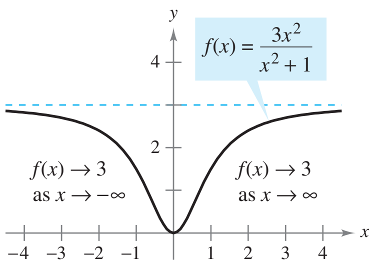
- **Determine (finite) limits at infinity.**
- **Determine the horizontal asymptotes, if any, of the graph of a function.**
- **Determine infinite limits at infinity.**

Limits at Infinity

This section discusses the "end behavior" of a function on an *infinite* interval. Consider the graph of

$$f(x) = \frac{3x^2}{x^2 + 1}$$

as shown in Figure 3.33. Graphically, you can see that the values of $f(x)$ appear to approach 3 as x increases without bound or decreases without bound. You can come to the same conclusions numerically, as shown in the table.



The limit of $f(x)$ as x approaches $-\infty$ or ∞ is 3.
Figure 3.33

← x decreases without bound. | x increases without bound. →

x	$-\infty \leftarrow$	-100	-10	-1	0	1	10	100	$\to \infty$
$f(x)$	$3 \leftarrow$	2.9997	2.97	1.5	0	1.5	2.97	2.9997	$\to 3$

← $f(x)$ approaches 3. | $f(x)$ approaches 3. →

The table suggests that the value of $f(x)$ approaches 3 as x increases without bound $(x \to \infty)$. Similarly, $f(x)$ approaches 3 as x decreases without bound $(x \to -\infty)$. These **limits at infinity** are denoted by

$$\lim_{x \to -\infty} f(x) = 3 \qquad \text{Limit at negative infinity}$$

and

$$\lim_{x \to \infty} f(x) = 3. \qquad \text{Limit at positive infinity}$$

NOTE The statement $\lim_{x \to -\infty} f(x) = L$ or $\lim_{x \to \infty} f(x) = L$, means that the limit exists *and* the limit is equal to L.

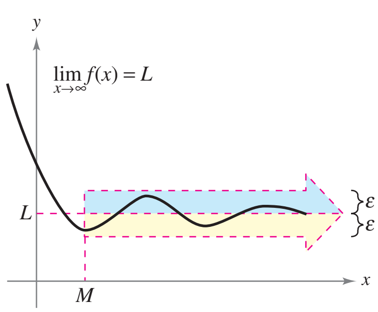
To say that a statement is true as x increases *without bound* means that for some (large) real number M, the statement is true for *all* x in the interval $\{x: x > M\}$. The following definition uses this concept.

Definition of Limits at Infinity

Let L be a real number.

1. The statement $\lim_{x \to \infty} f(x) = L$ means that for each $\varepsilon > 0$ there exists an $M > 0$ such that $|f(x) - L| < \varepsilon$ whenever $x > M$.

2. The statement $\lim_{x \to -\infty} f(x) = L$ means that for each $\varepsilon > 0$ there exists an $N < 0$ such that $|f(x) - L| < \varepsilon$ whenever $x < N$.



$f(x)$ is within ε units of L as $x \to \infty$.
Figure 3.34

The definition of a limit at infinity is shown in Figure 3.34. In this figure, note that for a given positive number ε there exists a positive number M such that, for $x > M$, the graph of f will lie between the horizontal lines given by $y = L + \varepsilon$ and $y = L - \varepsilon$.

Video Video

Horizontal Asymptotes

In Figure 3.34, the graph of f approaches the line $y = L$ as x increases without bound. The line $y = L$ is called a **horizontal asymptote** of the graph of f.

Definition of a Horizontal Asymptote

The line $y = L$ is a **horizontal asymptote** of the graph of f if

$$\lim_{x\to-\infty} f(x) = L \quad \text{or} \quad \lim_{x\to\infty} f(x) = L.$$

Note that from this definition, it follows that the graph of a *function* of x can have at most two horizontal asymptotes—one to the right and one to the left.

Limits at infinity have many of the same properties of limits discussed in Section 1.3. For example, if $\lim_{x\to\infty} f(x)$ and $\lim_{x\to\infty} g(x)$ both exist, then

$$\lim_{x\to\infty} [f(x) + g(x)] = \lim_{x\to\infty} f(x) + \lim_{x\to\infty} g(x)$$

and

$$\lim_{x\to\infty} [f(x)g(x)] = \left[\lim_{x\to\infty} f(x)\right]\left[\lim_{x\to\infty} g(x)\right].$$

Similar properties hold for limits at $-\infty$.

When evaluating limits at infinity, the following theorem is helpful. (A proof of this theorem is given in Appendix A.)

THEOREM 3.10 Limits at Infinity

If r is a positive rational number and c is any real number, then

$$\lim_{x\to\infty} \frac{c}{x^r} = 0.$$

Furthermore, if x^r is defined when $x < 0$, then

$$\lim_{x\to-\infty} \frac{c}{x^r} = 0.$$

EXAMPLE 1 Finding a Limit at Infinity

Find the limit: $\lim_{x\to\infty} \left(5 - \frac{2}{x^2}\right)$.

Solution Using Theorem 3.10, you can write

$$\begin{aligned} \lim_{x\to\infty} \left(5 - \frac{2}{x^2}\right) &= \lim_{x\to\infty} 5 - \lim_{x\to\infty} \frac{2}{x^2} && \text{Property of limits} \\ &= 5 - 0 \\ &= 5. \end{aligned}$$

Try It **Exploration A**

EXPLORATION

Use a graphing utility to graph

$$f(x) = \frac{2x^2 + 4x - 6}{3x^2 + 2x - 16}.$$

Describe all the important features of the graph. Can you find a single viewing window that shows all of these features clearly? Explain your reasoning.

What are the horizontal asymptotes of the graph? How far to the right do you have to move on the graph so that the graph is within 0.001 unit of its horizontal asymptote? Explain your reasoning.

EXAMPLE 2 Finding a Limit at Infinity

Find the limit: $\displaystyle\lim_{x\to\infty} \frac{2x-1}{x+1}$.

Solution Note that both the numerator and the denominator approach infinity as x approaches infinity.

$$\lim_{x\to\infty} \frac{2x-1}{x+1} \qquad \lim_{x\to\infty} (2x-1) \to \infty$$

$$\lim_{x\to\infty} (x+1) \to \infty$$

NOTE When you encounter an indeterminate form such as the one in Example 2, you should divide the numerator and denominator by the highest power of x in the *denominator*.

This results in $\frac{\infty}{\infty}$, an **indeterminate form.** To resolve this problem, you can divide both the numerator and the denominator by x. After dividing, the limit may be evaluated as shown.

$$\lim_{x\to\infty} \frac{2x-1}{x+1} = \lim_{x\to\infty} \frac{\frac{2x-1}{x}}{\frac{x+1}{x}} \qquad \text{Divide numerator and denominator by } x.$$

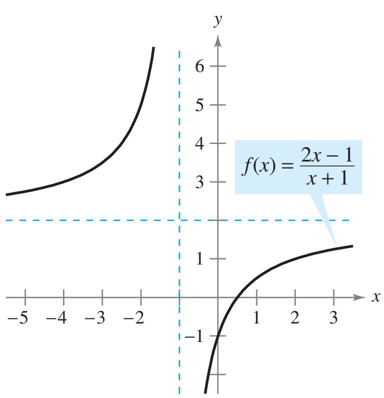
$$= \lim_{x\to\infty} \frac{2-\frac{1}{x}}{1+\frac{1}{x}} \qquad \text{Simplify.}$$

$$= \frac{\lim\limits_{x\to\infty} 2 - \lim\limits_{x\to\infty} \frac{1}{x}}{\lim\limits_{x\to\infty} 1 + \lim\limits_{x\to\infty} \frac{1}{x}} \qquad \text{Take limits of numerator and denominator.}$$

$$= \frac{2-0}{1+0} \qquad \text{Apply Theorem 3.10.}$$

$$= 2$$

So, the line $y = 2$ is a horizontal asymptote to the right. By taking the limit as $x \to -\infty$, you can see that $y = 2$ is also a horizontal asymptote to the left. The graph of the function is shown in Figure 3.35.



$y = 2$ is a horizontal asymptote.
Figure 3.35

Editable Graph

Try It   Exploration A   Exploration B

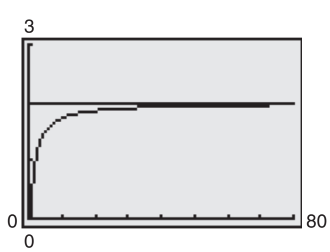
TECHNOLOGY You can test the reasonableness of the limit found in Example 2 by evaluating $f(x)$ for a few large positive values of x. For instance,

$$f(100) \approx 1.9703, \quad f(1000) \approx 1.9970, \quad \text{and} \quad f(10{,}000) \approx 1.9997.$$

Another way to test the reasonableness of the limit is to use a graphing utility. For instance, in Figure 3.36, the graph of

$$f(x) = \frac{2x-1}{x+1}$$

is shown with the horizontal line $y = 2$. Note that as x increases, the graph of f moves closer and closer to its horizontal asymptote.



As x increases, the graph of f moves closer and closer to the line $y = 2$.
Figure 3.36

EXAMPLE 3 A Comparison of Three Rational Functions

Find each limit.

a. $\lim_{x\to\infty} \dfrac{2x + 5}{3x^2 + 1}$ **b.** $\lim_{x\to\infty} \dfrac{2x^2 + 5}{3x^2 + 1}$ **c.** $\lim_{x\to\infty} \dfrac{2x^3 + 5}{3x^2 + 1}$

Solution In each case, attempting to evaluate the limit produces the indeterminate form ∞/∞.

a. Divide both the numerator and the denominator by x^2.

$$\lim_{x\to\infty} \frac{2x + 5}{3x^2 + 1} = \lim_{x\to\infty} \frac{(2/x) + (5/x^2)}{3 + (1/x^2)} = \frac{0 + 0}{3 + 0} = \frac{0}{3} = 0$$

b. Divide both the numerator and the denominator by x^2.

$$\lim_{x\to\infty} \frac{2x^2 + 5}{3x^2 + 1} = \lim_{x\to\infty} \frac{2 + (5/x^2)}{3 + (1/x^2)} = \frac{2 + 0}{3 + 0} = \frac{2}{3}$$

c. Divide both the numerator and the denominator by x^2.

$$\lim_{x\to\infty} \frac{2x^3 + 5}{3x^2 + 1} = \lim_{x\to\infty} \frac{2x + (5/x^2)}{3 + (1/x^2)} = \frac{\infty}{3}$$

You can conclude that the limit *does not exist* because the numerator increases without bound while the denominator approaches 3.

Try It Exploration A Exploration B

Open Exploration Video

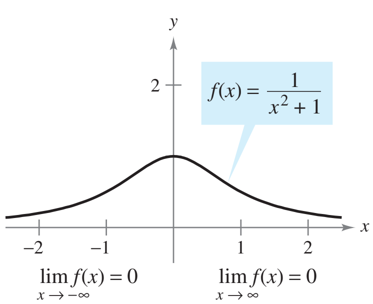
MARIA AGNESI (1718–1799)

Agnesi was one of a handful of women to receive credit for significant contributions to mathematics before the twentieth century. In her early twenties, she wrote the first text that included both differential and integral calculus. By age 30, she was an honorary member of the faculty at the University of Bologna.

MathBio

Guidelines for Finding Limits at $\pm\infty$ of Rational Functions

1. If the degree of the numerator is *less than* the degree of the denominator, then the limit of the rational function is 0.
2. If the degree of the numerator is *equal to* the degree of the denominator, then the limit of the rational function is the ratio of the leading coefficients.
3. If the degree of the numerator is *greater than* the degree of the denominator, then the limit of the rational function does not exist.



f has a horizontal asymptote at $y = 0$.
Figure 3.37

Use these guidelines to check the results in Example 3. These limits seem reasonable when you consider that for large values of x, the highest-power term of the rational function is the most "influential" in determining the limit. For instance, the limit as x approaches infinity of the function

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

is 0 because the denominator overpowers the numerator as x increases or decreases without bound, as shown in Figure 3.37.

The function shown in Figure 3.37 is a special case of a type of curve studied by the Italian mathematician Maria Gaetana Agnesi. The general form of this function is

$$f(x) = \frac{8a^3}{x^2 + 4a^2} \qquad \text{Witch of Agnesi}$$

and, through a mistranslation of the Italian word *vertéré*, the curve has come to be known as the Witch of Agnesi. Agnesi's work with this curve first appeared in a comprehensive text on calculus that was published in 1748.

FOR FURTHER INFORMATION For more information on the contributions of women to mathematics, see the article "Why Women Succeed in Mathematics" by Mona Fabricant, Sylvia Svitak, and Patricia Clark Kenschaft in *Mathematics Teacher*.

MathArticle

In Figure 3.37, you can see that the function $f(x) = 1/(x^2 + 1)$ approaches the same horizontal asymptote to the right and to the left. This is always true of rational functions. Functions that are not rational, however, may approach different horizontal asymptotes to the right and to the left. This is demonstrated in Example 4.

EXAMPLE 4 A Function with Two Horizontal Asymptotes

Find each limit.

a. $\displaystyle\lim_{x\to\infty} \frac{3x - 2}{\sqrt{2x^2 + 1}}$ **b.** $\displaystyle\lim_{x\to-\infty} \frac{3x - 2}{\sqrt{2x^2 + 1}}$

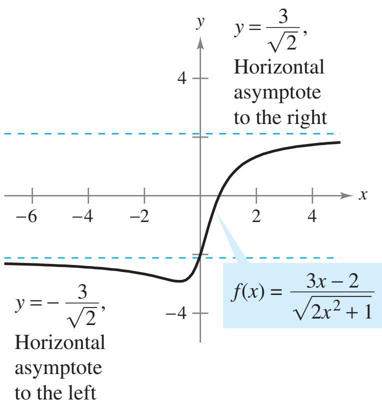
Solution

a. For $x > 0$, you can write $x = \sqrt{x^2}$. So, dividing both the numerator and the denominator by x produces

$$\frac{3x - 2}{\sqrt{2x^2 + 1}} = \frac{\dfrac{3x - 2}{x}}{\dfrac{\sqrt{2x^2 + 1}}{\sqrt{x^2}}} = \frac{3 - \dfrac{2}{x}}{\sqrt{\dfrac{2x^2 + 1}{x^2}}} = \frac{3 - \dfrac{2}{x}}{\sqrt{2 + \dfrac{1}{x^2}}}$$

and you can take the limit as follows.

$$\lim_{x\to\infty} \frac{3x - 2}{\sqrt{2x^2 + 1}} = \lim_{x\to\infty} \frac{3 - \dfrac{2}{x}}{\sqrt{2 + \dfrac{1}{x^2}}} = \frac{3 - 0}{\sqrt{2 + 0}} = \frac{3}{\sqrt{2}}$$



Functions that are not rational may have different right and left horizontal asymptotes.
Figure 3.38

Editable Graph

b. For $x < 0$, you can write $x = -\sqrt{x^2}$. So, dividing both the numerator and the denominator by x produces

$$\frac{3x - 2}{\sqrt{2x^2 + 1}} = \frac{\dfrac{3x - 2}{x}}{\dfrac{\sqrt{2x^2 + 1}}{-\sqrt{x^2}}} = \frac{3 - \dfrac{2}{x}}{-\sqrt{\dfrac{2x^2 + 1}{x^2}}} = \frac{3 - \dfrac{2}{x}}{-\sqrt{2 + \dfrac{1}{x^2}}}$$

and you can take the limit as follows.

$$\lim_{x\to-\infty} \frac{3x - 2}{\sqrt{2x^2 + 1}} = \lim_{x\to-\infty} \frac{3 - \dfrac{2}{x}}{-\sqrt{2 + \dfrac{1}{x^2}}} = \frac{3 - 0}{-\sqrt{2 + 0}} = -\frac{3}{\sqrt{2}}$$

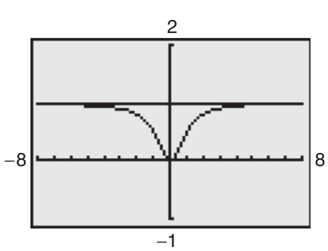
The graph of $f(x) = (3x - 2)/\sqrt{2x^2 + 1}$ is shown in Figure 3.38.

Try It Exploration A

TECHNOLOGY PITFALL If you use a graphing utility to help estimate a limit, be sure that you also confirm the estimate analytically—the pictures shown by a graphing utility can be misleading. For instance, Figure 3.39 shows one view of the graph of

$$y = \frac{2x^3 + 1000x^2 + x}{x^3 + 1000x^2 + x + 1000}.$$

From this view, one could be convinced that the graph has $y = 1$ as a horizontal asymptote. An analytical approach shows that the horizontal asymptote is actually $y = 2$. Confirm this by enlarging the viewing window on the graphing utility.



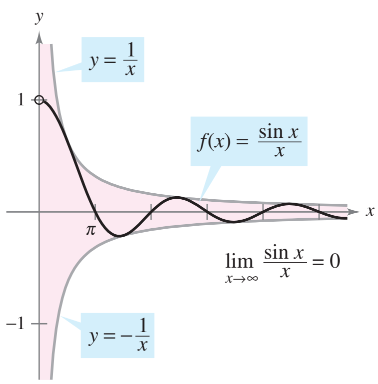
The horizontal asymptote appears to be the line $y = 1$ but it is actually the line $y = 2$.
Figure 3.39

In Section 1.3 (Example 9), you saw how the Squeeze Theorem can be used to evaluate limits involving trigonometric functions. This theorem is also valid for limits at infinity.

EXAMPLE 5 Limits Involving Trigonometric Functions

Find each limit.

a. $\lim_{x\to\infty} \sin x$ **b.** $\lim_{x\to\infty} \dfrac{\sin x}{x}$



As x increases without bound, $f(x)$ approaches 0.
Figure 3.40

Editable Graph

Solution

a. As x approaches infinity, the sine function oscillates between 1 and -1. So, this limit does not exist.

b. Because $-1 \le \sin x \le 1$, it follows that for $x > 0$,

$$-\frac{1}{x} \le \frac{\sin x}{x} \le \frac{1}{x}$$

where $\lim_{x\to\infty} (-1/x) = 0$ and $\lim_{x\to\infty} (1/x) = 0$. So, by the Squeeze Theorem, you can obtain

$$\lim_{x\to\infty} \frac{\sin x}{x} = 0$$

as shown in Figure 3.40.

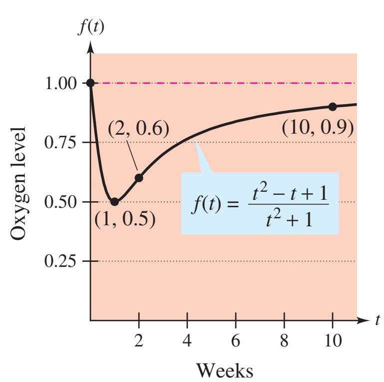
Try It Exploration A

EXAMPLE 6 Oxygen Level in a Pond

Suppose that $f(t)$ measures the level of oxygen in a pond, where $f(t) = 1$ is the normal (unpolluted) level and the time t is measured in weeks. When $t = 0$, organic waste is dumped into the pond, and as the waste material oxidizes, the level of oxygen in the pond is

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

What percent of the normal level of oxygen exists in the pond after 1 week? After 2 weeks? After 10 weeks? What is the limit as t approaches infinity?



The level of oxygen in a pond approaches the normal level of 1 as t approaches ∞.
Figure 3.41

Solution When $t = 1$, 2, and 10, the levels of oxygen are as shown.

$$f(1) = \frac{1^2 - 1 + 1}{1^2 + 1} = \frac{1}{2} = 50\% \qquad \text{1 week}$$

$$f(2) = \frac{2^2 - 2 + 1}{2^2 + 1} = \frac{3}{5} = 60\% \qquad \text{2 weeks}$$

$$f(10) = \frac{10^2 - 10 + 1}{10^2 + 1} = \frac{91}{101} \approx 90.1\% \qquad \text{10 weeks}$$

To find the limit as t approaches infinity, divide the numerator and the denominator by t^2 to obtain

$$\lim_{t\to\infty} \frac{t^2 - t + 1}{t^2 + 1} = \lim_{t\to\infty} \frac{1 - (1/t) + (1/t^2)}{1 + (1/t^2)} = \frac{1 - 0 + 0}{1 + 0} = 1 = 100\%.$$

See Figure 3.41.

Try It Exploration A

Infinite Limits at Infinity

Many functions do not approach a finite limit as x increases (or decreases) without bound. For instance, no polynomial function has a finite limit at infinity. The following definition is used to describe the behavior of polynomial and other functions at infinity.

NOTE Determining whether a function has an infinite limit at infinity is useful in analyzing the "end behavior" of its graph. You will see examples of this in Section 3.6 on curve sketching.

Definition of Infinite Limits at Infinity

Let f be a function defined on the interval (a, ∞).

1. The statement $\lim_{x\to\infty} f(x) = \infty$ means that for each positive number M, there is a corresponding number $N > 0$ such that $f(x) > M$ whenever $x > N$.
2. The statement $\lim_{x\to\infty} f(x) = -\infty$ means that for each negative number M, there is a corresponding number $N > 0$ such that $f(x) < M$ whenever $x > N$.

Similar definitions can be given for the statements $\lim_{x\to-\infty} f(x) = \infty$ and $\lim_{x\to-\infty} f(x) = -\infty$.

EXAMPLE 7 Finding Infinite Limits at Infinity

Find each limit.

a. $\lim_{x\to\infty} x^3$ **b.** $\lim_{x\to-\infty} x^3$

Solution

a. As x increases without bound, x^3 also increases without bound. So, you can write $\lim_{x\to\infty} x^3 = \infty$.

b. As x decreases without bound, x^3 also decreases without bound. So, you can write $\lim_{x\to-\infty} x^3 = -\infty$.

The graph of $f(x) = x^3$ in Figure 3.42 illustrates these two results. These results agree with the Leading Coefficient Test for polynomial functions as described in Section P.3.

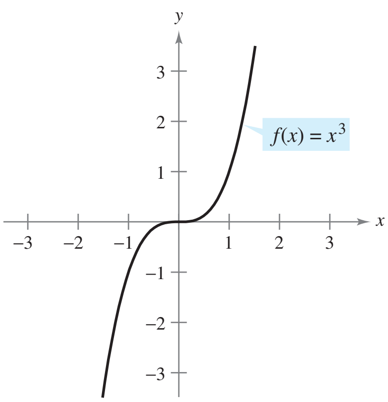


Figure 3.42

Editable Graph

Try It Exploration A

EXAMPLE 8 Finding Infinite Limits at Infinity

Find each limit.

a. $\lim_{x\to\infty} \dfrac{2x^2 - 4x}{x + 1}$ **b.** $\lim_{x\to-\infty} \dfrac{2x^2 - 4x}{x + 1}$

Solution One way to evaluate each of these limits is to use long division to rewrite the improper rational function as the sum of a polynomial and a rational function.

$$\textbf{a.}\quad \lim_{x\to\infty} \frac{2x^2 - 4x}{x + 1} = \lim_{x\to\infty}\left(2x - 6 + \frac{6}{x + 1}\right) = \infty$$

$$\textbf{b.}\quad \lim_{x\to-\infty} \frac{2x^2 - 4x}{x + 1} = \lim_{x\to-\infty}\left(2x - 6 + \frac{6}{x + 1}\right) = -\infty$$

The statements above can be interpreted as saying that as x approaches $\pm\infty$, the function $f(x) = (2x^2 - 4x)/(x + 1)$ behaves like the function $g(x) = 2x - 6$. In Section 3.6, you will see that this is graphically described by saying that the line $y = 2x - 6$ is a slant asymptote of the graph of f, as shown in Figure 3.43.

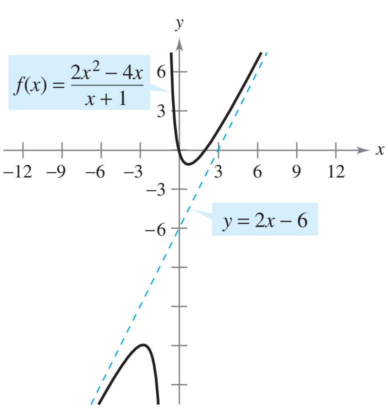


Figure 3.43

Editable Graph

Try It Exploration A

Section 3.6 A Summary of Curve Sketching

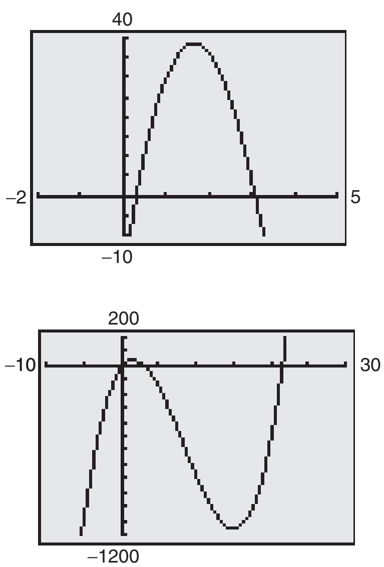
- **Analyze and sketch the graph of a function.**

Analyzing the Graph of a Function

It would be difficult to overstate the importance of using graphs in mathematics. Descartes's introduction of analytic geometry contributed significantly to the rapid advances in calculus that began during the mid-seventeenth century. In the words of Lagrange, "As long as algebra and geometry traveled separate paths their advance was slow and their applications limited. But when these two sciences joined company, they drew from each other fresh vitality and thenceforth marched on at a rapid pace toward perfection."

So far, you have studied several concepts that are useful in analyzing the graph of a function.

- x-intercepts and y-intercepts (Section P.1)
- Symmetry (Section P.1)
- Domain and range (Section P.3)
- Continuity (Section 1.4)
- Vertical asymptotes (Section 1.5)
- Differentiability (Section 2.1)
- Relative extrema (Section 3.1)
- Concavity (Section 3.4)
- Points of inflection (Section 3.4)
- Horizontal asymptotes (Section 3.5)
- Infinite limits at infinity (Section 3.5)



Different viewing windows for the graph of $f(x) = x^3 - 25x^2 + 74x - 20$
Figure 3.44

When you are sketching the graph of a function, either by hand or with a graphing utility, remember that normally you cannot show the *entire* graph. The decision as to which part of the graph you choose to show is often crucial. For instance, which of the viewing windows in Figure 3.44 better represents the graph of

$$f(x) = x^3 - 25x^2 + 74x - 20?$$

By seeing both views, it is clear that the second viewing window gives a more complete representation of the graph. But would a third viewing window reveal other interesting portions of the graph? To answer this, you need to use calculus to interpret the first and second derivatives. Here are some guidelines for determining a good viewing window for the graph of a function.

Guidelines for Analyzing the Graph of a Function

1. Determine the domain and range of the function.
2. Determine the intercepts, asymptotes, and symmetry of the graph.
3. Locate the x-values for which $f'(x)$ and $f''(x)$ either are zero or do not exist. Use the results to determine relative extrema and points of inflection.

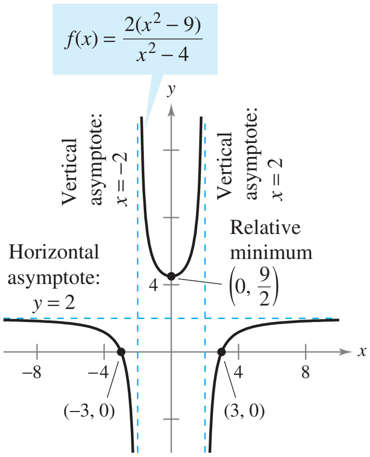
NOTE In these guidelines, note the importance of *algebra* (as well as calculus) for solving the equations $f(x) = 0$, $f'(x) = 0$, and $f''(x) = 0$.

EXAMPLE 1 Sketching the Graph of a Rational Function

Analyze and sketch the graph of $f(x) = \dfrac{2(x^2 - 9)}{x^2 - 4}$.

Solution

First derivative:	$f'(x) = \dfrac{20x}{(x^2 - 4)^2}$
Second derivative:	$f''(x) = \dfrac{-20(3x^2 + 4)}{(x^2 - 4)^3}$
x-intercepts:	$(-3, 0), (3, 0)$
y-intercept:	$\left(0, \frac{9}{2}\right)$
Vertical asymptotes:	$x = -2, x = 2$
Horizontal asymptote:	$y = 2$
Critical number:	$x = 0$
Possible points of inflection:	None
Domain:	All real numbers except $x = \pm 2$
Symmetry:	With respect to y-axis
Test intervals:	$(-\infty, -2), (-2, 0), (0, 2), (2, \infty)$



Using calculus, you can be certain that you have determined all characteristics of the graph of f.

Figure 3.45

Editable Graph

The table shows how the test intervals are used to determine several characteristics of the graph. The graph of f is shown in Figure 3.45.

	$f(x)$	$f'(x)$	$f''(x)$	Characteristic of Graph
$-\infty < x < -2$		$-$	$-$	Decreasing, concave downward
$x = -2$	Undef.	Undef.	Undef.	Vertical asymptote
$-2 < x < 0$		$-$	$+$	Decreasing, concave upward
$x = 0$	$\frac{9}{2}$	0	$+$	Relative minimum
$0 < x < 2$		$+$	$+$	Increasing, concave upward
$x = 2$	Undef.	Undef.	Undef.	Vertical asymptote
$2 < x < \infty$		$+$	$-$	Increasing, concave downward

FOR FURTHER INFORMATION For more information on the use of technology to graph rational functions, see the article "Graphs of Rational Functions for Computer Assisted Calculus" by Stan Byrd and Terry Walters in *The College Mathematics Journal*.

MathArticle

Try It   Exploration A

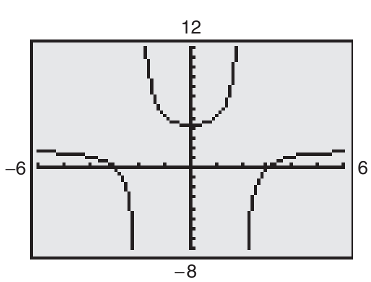
Video   Video   Video   Video

Be sure you understand all of the implications of creating a table such as that shown in Example 1. Because of the use of calculus, you can *be sure* that the graph has no relative extrema or points of inflection other than those shown in Figure 3.45.

TECHNOLOGY PITFALL Without using the type of analysis outlined in Example 1, it is easy to obtain an incomplete view of a graph's basic characteristics. For instance, Figure 3.46 shows a view of the graph of

$$g(x) = \frac{2(x^2 - 9)(x - 20)}{(x^2 - 4)(x - 21)}.$$

From this view, it appears that the graph of g is about the same as the graph of f shown in Figure 3.45. The graphs of these two functions, however, differ significantly. Try enlarging the viewing window to see the differences.



By not using calculus you may overlook important characteristics of the graph of g.

Figure 3.46

EXAMPLE 2 Sketching the Graph of a Rational Function

Analyze and sketch the graph of $f(x) = \dfrac{x^2 - 2x + 4}{x - 2}$.

Solution

First derivative: $f'(x) = \dfrac{x(x-4)}{(x-2)^2}$

Second derivative: $f''(x) = \dfrac{8}{(x-2)^3}$

x-intercepts: None

y-intercept: $(0, -2)$

Vertical asymptote: $x = 2$

Horizontal asymptotes: None

End behavior: $\lim_{x\to-\infty} f(x) = -\infty, \ \lim_{x\to\infty} f(x) = \infty$

Critical numbers: $x = 0, x = 4$

Possible points of inflection: None

Domain: All real numbers except $x = 2$

Test intervals: $(-\infty, 0), (0, 2), (2, 4), (4, \infty)$

The analysis of the graph of f is shown in the table, and the graph is shown in Figure 3.47.

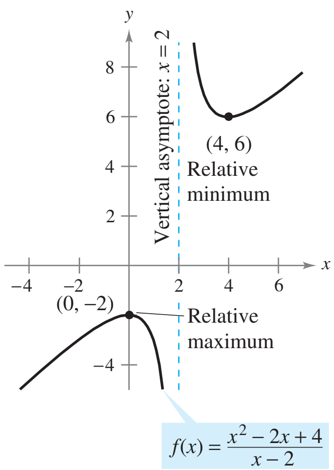


Figure 3.47

Editable Graph

	$f(x)$	$f'(x)$	$f''(x)$	Characteristic of Graph
$-\infty < x < 0$		$+$	$-$	Increasing, concave downward
$x = 0$	-2	0	$-$	Relative maximum
$0 < x < 2$		$-$	$-$	Decreasing, concave downward
$x = 2$	Undef.	Undef.	Undef.	Vertical asymptote
$2 < x < 4$		$-$	$+$	Decreasing, concave upward
$x = 4$	6	0	$+$	Relative minimum
$4 < x < \infty$		$+$	$+$	Increasing, concave upward

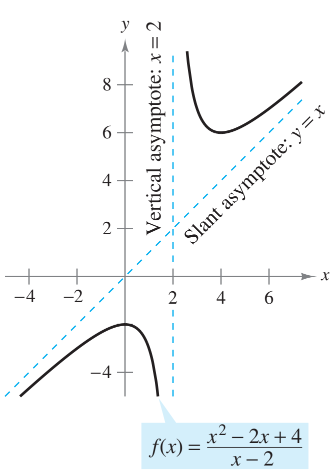
Try It Exploration A

Although the graph of the function in Example 2 has no horizontal asymptote, it does have a slant asymptote. The graph of a rational function (having no common factors and whose denominator is of degree 1 or greater) has a **slant asymptote** if the degree of the numerator exceeds the degree of the denominator by exactly 1. To find the slant asymptote, use long division to rewrite the rational function as the sum of a first-degree polynomial and another rational function.

$$f(x) = \frac{x^2 - 2x + 4}{x - 2} \qquad \text{Write original equation.}$$

$$= x + \frac{4}{x - 2} \qquad \text{Rewrite using long division.}$$

In Figure 3.48, note that the graph of f approaches the slant asymptote $y = x$ as x approaches $-\infty$ or ∞.



A slant asymptote
Figure 3.48

EXAMPLE 3 Sketching the Graph of a Radical Function

Analyze and sketch the graph of $f(x) = \dfrac{x}{\sqrt{x^2 + 2}}$.

Solution

$$f'(x) = \frac{2}{(x^2 + 2)^{3/2}} \qquad f''(x) = -\frac{6x}{(x^2 + 2)^{5/2}}$$

The graph has only one intercept, $(0, 0)$. It has no vertical asymptotes, but it has two horizontal asymptotes: $y = 1$ (to the right) and $y = -1$ (to the left). The function has no critical numbers and one possible point of inflection (at $x = 0$). The domain of the function is all real numbers, and the graph is symmetric with respect to the origin. The analysis of the graph of f is shown in the table, and the graph is shown in Figure 3.49.

	$f(x)$	$f'(x)$	$f''(x)$	Characteristic of Graph
$-\infty < x < 0$		$+$	$+$	Increasing, concave upward
$x = 0$	0	$\dfrac{1}{\sqrt{2}}$	0	Point of inflection
$0 < x < \infty$		$+$	$-$	Increasing, concave downward

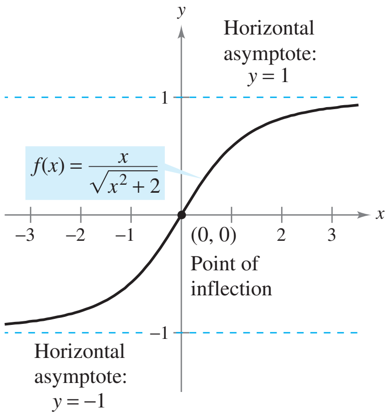


Figure 3.49

Editable Graph

Try It Exploration A

EXAMPLE 4 Sketching the Graph of a Radical Function

Analyze and sketch the graph of $f(x) = 2x^{5/3} - 5x^{4/3}$.

Solution

$$f'(x) = \frac{10}{3}x^{1/3}(x^{1/3} - 2) \qquad f''(x) = \frac{20(x^{1/3} - 1)}{9x^{2/3}}$$

The function has two intercepts: $(0, 0)$ and $\left(\frac{125}{8}, 0\right)$. There are no horizontal or vertical asymptotes. The function has two critical numbers ($x = 0$ and $x = 8$) and two possible points of inflection ($x = 0$ and $x = 1$). The domain is all real numbers. The analysis of the graph of f is shown in the table, and the graph is shown in Figure 3.50.

	$f(x)$	$f'(x)$	$f''(x)$	Characteristic of Graph
$-\infty < x < 0$		$+$	$-$	Increasing, concave downward
$x = 0$	0	0	Undef.	Relative maximum
$0 < x < 1$		$-$	$-$	Decreasing, concave downward
$x = 1$	-3	$-$	0	Point of inflection
$1 < x < 8$		$-$	$+$	Decreasing, concave upward
$x = 8$	-16	0	$+$	Relative minimum
$8 < x < \infty$		$+$	$+$	Increasing, concave upward

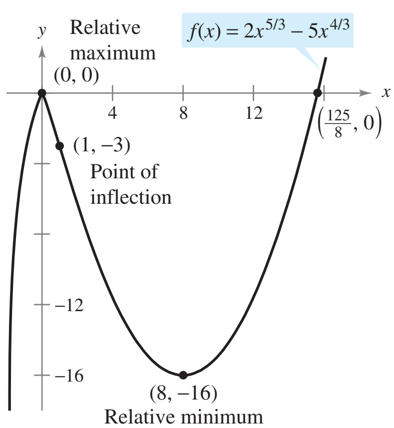


Figure 3.50

Editable Graph

Try It Exploration A

EXAMPLE 5 Sketching the Graph of a Polynomial Function

Analyze and sketch the graph of $f(x) = x^4 - 12x^3 + 48x^2 - 64x$.

Solution Begin by factoring to obtain

$$\begin{aligned} f(x) &= x^4 - 12x^3 + 48x^2 - 64x \\ &= x(x-4)^3. \end{aligned}$$

Then, using the factored form of $f(x)$, you can perform the following analysis.

First derivative: $f'(x) = 4(x-1)(x-4)^2$

Second derivative: $f''(x) = 12(x-4)(x-2)$

x-intercepts: $(0, 0), (4, 0)$

y-intercept: $(0, 0)$

Vertical asymptotes: None

Horizontal asymptotes: None

End behavior: $\lim_{x\to-\infty} f(x) = \infty, \lim_{x\to\infty} f(x) = \infty$

Critical numbers: $x = 1, x = 4$

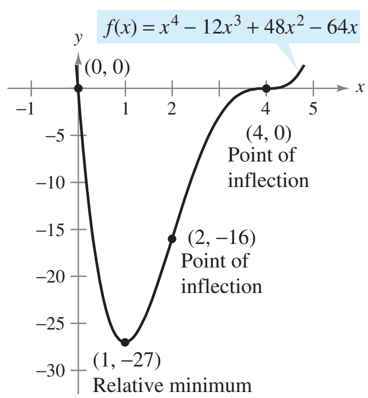
Possible points of inflection: $x = 2, x = 4$

Domain: All real numbers

Test intervals: $(-\infty, 1), (1, 2), (2, 4), (4, \infty)$

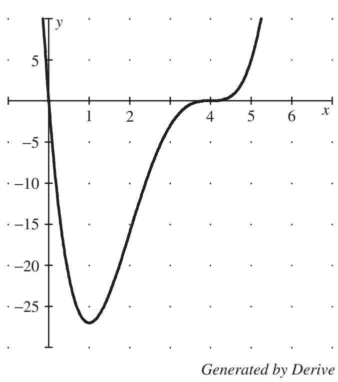
The analysis of the graph of f is shown in the table, and the graph is shown in Figure 3.51(a). Using a computer algebra system such as *Derive* [see Figure 3.51(b)] can help you verify your analysis.

	$f(x)$	$f'(x)$	$f''(x)$	Characteristic of Graph
$-\infty < x < 1$		$-$	$+$	Decreasing, concave upward
$x = 1$	-27	0	$+$	Relative minimum
$1 < x < 2$		$+$	$+$	Increasing, concave upward
$x = 2$	-16	$+$	0	Point of inflection
$2 < x < 4$		$+$	$-$	Increasing, concave downward
$x = 4$	0	0	0	Point of inflection
$4 < x < \infty$		$+$	$+$	Increasing, concave upward



(a)

Editable Graph



(b)

A polynomial function of even degree must have at least one relative extremum.

Figure 3.51

Try It | Exploration A | Exploration B | Open Exploration

The fourth-degree polynomial function in Example 5 has one relative minimum and no relative maxima. In general, a polynomial function of degree n can have *at most* $n - 1$ relative extrema, and *at most* $n - 2$ points of inflection. Moreover, polynomial functions of even degree must have *at least* one relative extremum.

Remember from the Leading Coefficient Test described in Section P.3 that the "end behavior" of the graph of a polynomial function is determined by its leading coefficient and its degree. For instance, because the polynomial in Example 5 has a positive leading coefficient, the graph rises to the right. Moreover, because the degree is even, the graph also rises to the left.

EXAMPLE 6 Sketching the Graph of a Trigonometric Function

Analyze and sketch the graph of $f(x) = \dfrac{\cos x}{1 + \sin x}$.

Solution Because the function has a period of 2π, you can restrict the analysis of the graph to any interval of length 2π. For convenience, choose $(-\pi/2, 3\pi/2)$.

First derivative: $f'(x) = -\dfrac{1}{1 + \sin x}$

Second derivative: $f''(x) = \dfrac{\cos x}{(1 + \sin x)^2}$

Period: 2π

x-intercept: $\left(\dfrac{\pi}{2}, 0\right)$

y-intercept: $(0, 1)$

Vertical asymptotes: $x = -\dfrac{\pi}{2},\ x = \dfrac{3\pi}{2}$ See Note below.

Horizontal asymptotes: None

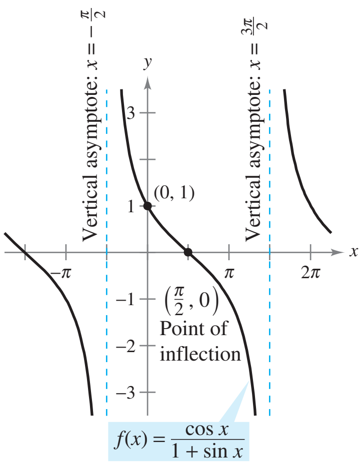
Critical numbers: None

Possible points of inflection: $x = \dfrac{\pi}{2}$

Domain: All real numbers except $x = \dfrac{3 + 4n}{2}\pi$

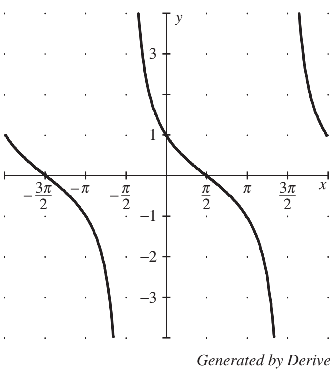
Test intervals: $\left(-\dfrac{\pi}{2}, \dfrac{\pi}{2}\right), \left(\dfrac{\pi}{2}, \dfrac{3\pi}{2}\right)$

The analysis of the graph of f on the interval $(-\pi/2, 3\pi/2)$ is shown in the table, and the graph is shown in Figure 3.52(a). Compare this with the graph generated by the computer algebra system *Derive* in Figure 3.52(b).



(a)

Editable Graph



(b)

Figure 3.52

	$f(x)$	$f'(x)$	$f''(x)$	Characteristic of Graph
$x = -\dfrac{\pi}{2}$	Undef.	Undef.	Undef.	Vertical asymptote
$-\dfrac{\pi}{2} < x < \dfrac{\pi}{2}$		$-$	$+$	Decreasing, concave upward
$x = \dfrac{\pi}{2}$	0	$-\dfrac{1}{2}$	0	Point of inflection
$\dfrac{\pi}{2} < x < \dfrac{3\pi}{2}$		$-$	$-$	Decreasing, concave downward
$x = \dfrac{3\pi}{2}$	Undef.	Undef.	Undef.	Vertical asymptote

Try It   Exploration A

NOTE By substituting $-\pi/2$ or $3\pi/2$ into the function, you obtain the form $0/0$. This is called an indeterminate form and you will study this in Section 8.7. To determine that the function has vertical asymptotes at these two values, you can rewrite the function as follows.

$$f(x) = \frac{\cos x}{1 + \sin x} = \frac{(\cos x)(1 - \sin x)}{(1 + \sin x)(1 - \sin x)} = \frac{(\cos x)(1 - \sin x)}{\cos^2 x} = \frac{1 - \sin x}{\cos x}$$

In this form, it is clear that the graph of f has vertical asymptotes when $x = -\pi/2$ and $3\pi/2$.

Section 3.7

Optimization Problems

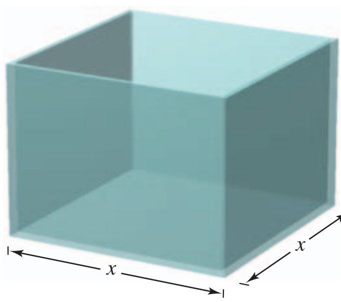
- **Solve applied minimum and maximum problems.**

Applied Minimum and Maximum Problems

One of the most common applications of calculus involves the determination of minimum and maximum values. Consider how frequently you hear or read terms such as greatest profit, least cost, least time, greatest voltage, optimum size, least size, greatest strength, and greatest distance. Before outlining a general problem-solving strategy for such problems, let's look at an example.

EXAMPLE 1 Finding Maximum Volume

A manufacturer wants to design an open box having a square base and a surface area of 108 square inches, as shown in Figure 3.53. What dimensions will produce a box with maximum volume?



Open box with square base:
$S = x^2 + 4xh = 108$
Figure 3.53

Solution Because the box has a square base, its volume is

$$V = x^2h. \qquad \text{Primary equation}$$

This equation is called the **primary equation** because it gives a formula for the quantity to be optimized. The surface area of the box is

$$S = (\text{area of base}) + (\text{area of four sides})$$
$$S = x^2 + 4xh = 108. \qquad \text{Secondary equation}$$

Because V is to be maximized, you want to write V as a function of just one variable. To do this, you can solve the equation $x^2 + 4xh = 108$ for h in terms of x to obtain $h = (108 - x^2)/(4x)$. Substituting into the primary equation produces

$$\begin{aligned} V &= x^2h && \text{Function of two variables} \\ &= x^2\left(\frac{108 - x^2}{4x}\right) && \text{Substitute for } h. \\ &= 27x - \frac{x^3}{4}. && \text{Function of one variable} \end{aligned}$$

Before finding which x-value will yield a maximum value of V, you should determine the *feasible domain*. That is, what values of x make sense in this problem? You know that $V \geq 0$. You also know that x must be nonnegative and that the area of the base $(A = x^2)$ is at most 108. So, the feasible domain is

$$0 \leq x \leq \sqrt{108}. \qquad \text{Feasible domain}$$

To maximize V, find the critical numbers of the volume function.

$$\begin{aligned} \frac{dV}{dx} = 27 - \frac{3x^2}{4} &= 0 && \text{Set derivative equal to 0.} \\ 3x^2 &= 108 && \text{Simplify.} \\ x &= \pm 6 && \text{Critical numbers} \end{aligned}$$

So, the critical numbers are $x = \pm 6$. You do not need to consider $x = -6$ because it is outside the domain. Evaluating V at the critical number $x = 6$ and at the endpoints of the domain produces $V(0) = 0$, $V(6) = 108$, and $V\left(\sqrt{108}\right) = 0$. So, V is maximum when $x = 6$ and the dimensions of the box are $6 \times 6 \times 3$ inches.

TECHNOLOGY You can verify your answer by using a graphing utility to graph the volume function

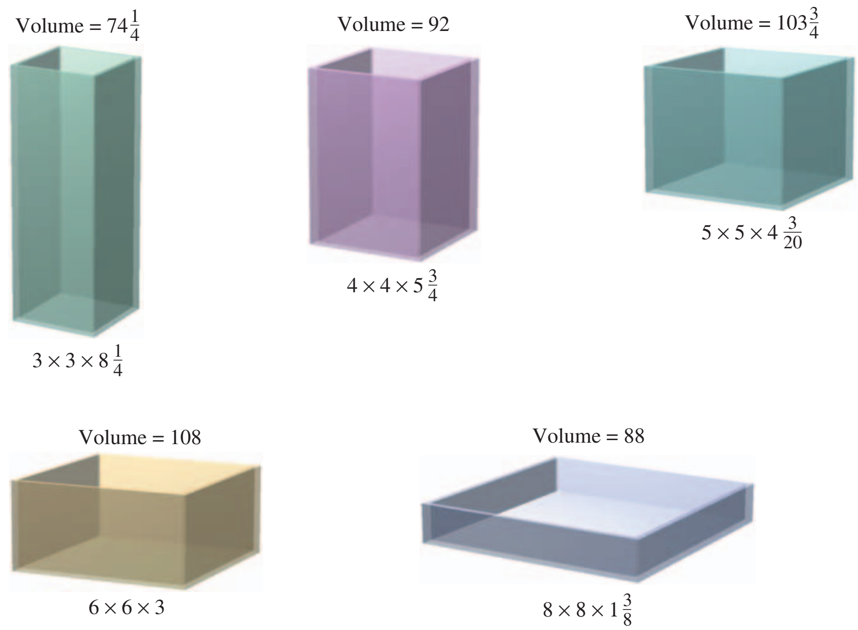
$$V = 27x - \frac{x^3}{4}.$$

Use a viewing window in which $0 \leq x \leq \sqrt{108} \approx 10.4$ and $0 \leq y \leq 120$, and the *trace* feature to determine the maximum value of V.

Try It | Exploration A | Video | Video

In Example 1, you should realize that there are infinitely many open boxes having 108 square inches of surface area. To begin solving the problem, you might ask yourself which basic shape would seem to yield a maximum volume. Should the box be tall, squat, or nearly cubical?

You might even try calculating a few volumes, as shown in Figure 3.54, to see if you can get a better feeling for what the optimum dimensions should be. Remember that you are not ready to begin solving a problem until you have clearly identified what the problem is.



Which box has the greatest volume?
Figure 3.54

Example 1 illustrates the following guidelines for solving applied minimum and maximum problems.

Guidelines for Solving Applied Minimum and Maximum Problems

1. Identify all *given* quantities and quantities *to be determined*. If possible, make a sketch.
2. Write a **primary equation** for the quantity that is to be maximized or minimized. (A review of several useful formulas from geometry is presented inside the front cover.)
3. Reduce the primary equation to one having a *single independent variable*. This may involve the use of **secondary equations** relating the independent variables of the primary equation.
4. Determine the feasible domain of the primary equation. That is, determine the values for which the stated problem makes sense.
5. Determine the desired maximum or minimum value by the calculus techniques discussed in Sections 3.1 through 3.4.

NOTE When performing Step 5, recall that to determine the maximum or minimum value of a continuous function f on a closed interval, you should compare the values of f at its critical numbers with the values of f at the endpoints of the interval.

EXAMPLE 2 Finding Minimum Distance

Which points on the graph of $y = 4 - x^2$ are closest to the point $(0, 2)$?

Solution Figure 3.55 shows that there are two points at a minimum distance from the point $(0, 2)$. The distance between the point $(0, 2)$ and a point (x, y) on the graph of $y = 4 - x^2$ is given by

$$d = \sqrt{(x - 0)^2 + (y - 2)^2}. \qquad \text{Primary equation}$$

Using the secondary equation $y = 4 - x^2$, you can rewrite the primary equation as

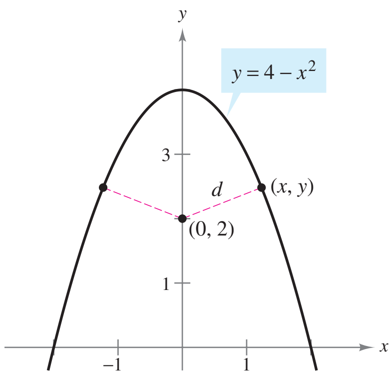
$$d = \sqrt{x^2 + (4 - x^2 - 2)^2} = \sqrt{x^4 - 3x^2 + 4}.$$

Because d is smallest when the expression inside the radical is smallest, you need only find the critical numbers of $f(x) = x^4 - 3x^2 + 4$. Note that the domain of f is the entire real line. So, there are no endpoints of the domain to consider. Moreover, setting $f'(x)$ equal to 0 yields

$$f'(x) = 4x^3 - 6x = 2x(2x^2 - 3) = 0$$

$$x = 0, \sqrt{\frac{3}{2}}, -\sqrt{\frac{3}{2}}.$$

The First Derivative Test verifies that $x = 0$ yields a relative maximum, whereas both $x = \sqrt{3/2}$ and $x = -\sqrt{3/2}$ yield a minimum distance. So, the closest points are $\left(\sqrt{3/2}, 5/2\right)$ and $\left(-\sqrt{3/2}, 5/2\right)$.



The quantity to be minimized is distance: $d = \sqrt{(x - 0)^2 + (y - 2)^2}$.
Figure 3.55

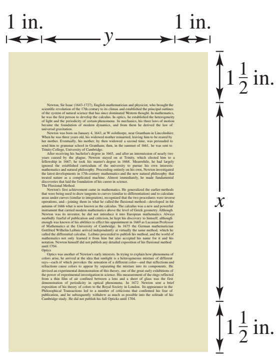
Editable Graph

Try It | Exploration A | Exploration B

Open Exploration | Video

EXAMPLE 3 Finding Minimum Area

A rectangular page is to contain 24 square inches of print. The margins at the top and bottom of the page are to be $1\frac{1}{2}$ inches, and the margins on the left and right are to be 1 inch (see Figure 3.56). What should the dimensions of the page be so that the least amount of paper is used?



The quantity to be minimized is area: $A = (x + 3)(y + 2)$.
Figure 3.56

Solution Let A be the area to be minimized.

$$A = (x + 3)(y + 2) \qquad \text{Primary equation}$$

The printed area inside the margins is given by

$$24 = xy. \qquad \text{Secondary equation}$$

Solving this equation for y produces $y = 24/x$. Substitution into the primary equation produces

$$A = (x + 3)\left(\frac{24}{x} + 2\right) = 30 + 2x + \frac{72}{x}. \qquad \text{Function of one variable}$$

Because x must be positive, you are interested only in values of A for $x > 0$. To find the critical numbers, differentiate with respect to x.

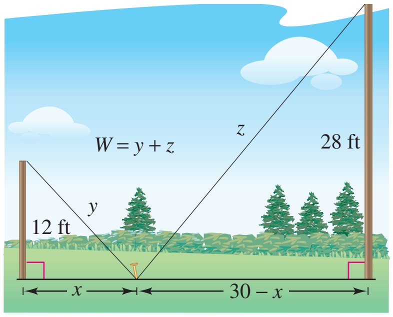
$$\frac{dA}{dx} = 2 - \frac{72}{x^2} = 0 \quad \Rightarrow \quad x^2 = 36$$

So, the critical numbers are $x = \pm 6$. You do not have to consider $x = -6$ because it is outside the domain. The First Derivative Test confirms that A is a minimum when $x = 6$. So, $y = \frac{24}{6} = 4$ and the dimensions of the page should be $x + 3 = 9$ inches by $y + 2 = 6$ inches.

Try It | Exploration A

EXAMPLE 4 Finding Minimum Length

Two posts, one 12 feet high and the other 28 feet high, stand 30 feet apart. They are to be stayed by two wires, attached to a single stake, running from ground level to the top of each post. Where should the stake be placed to use the least amount of wire?



The quantity to be minimized is length. From the diagram, you can see that x varies between 0 and 30.
Figure 3.57

Simulation

Solution Let W be the wire length to be minimized. Using Figure 3.57, you can write

$$W = y + z. \qquad \text{Primary equation}$$

In this problem, rather than solving for y in terms of z (or vice versa), you can solve for both y and z in terms of a third variable x, as shown in Figure 3.57. From the Pythagorean Theorem, you obtain

$$x^2 + 12^2 = y^2$$
$$(30 - x)^2 + 28^2 = z^2$$

which implies that

$$y = \sqrt{x^2 + 144}$$
$$z = \sqrt{x^2 - 60x + 1684}.$$

So, W is given by

$$\begin{aligned} W &= y + z \\ &= \sqrt{x^2 + 144} + \sqrt{x^2 - 60x + 1684}, \qquad 0 \le x \le 30. \end{aligned}$$

Differentiating W with respect to x yields

$$\frac{dW}{dx} = \frac{x}{\sqrt{x^2 + 144}} + \frac{x - 30}{\sqrt{x^2 - 60x + 1684}}.$$

By letting $dW/dx = 0$, you obtain

$$\begin{aligned} \frac{x}{\sqrt{x^2 + 144}} + \frac{x - 30}{\sqrt{x^2 - 60x + 1684}} &= 0 \\ x\sqrt{x^2 - 60x + 1684} &= (30 - x)\sqrt{x^2 + 144} \\ x^2(x^2 - 60x + 1684) &= (30 - x)^2(x^2 + 144) \\ x^4 - 60x^3 + 1684x^2 &= x^4 - 60x^3 + 1044x^2 - 8640x + 129{,}600 \\ 640x^2 + 8640x - 129{,}600 &= 0 \\ 320(x - 9)(2x + 45) &= 0 \\ x &= 9, -22.5. \end{aligned}$$

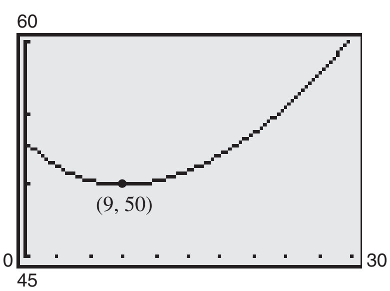
Because $x = -22.5$ is not in the domain and

$$W(0) \approx 53.04, \qquad W(9) = 50, \qquad \text{and} \qquad W(30) \approx 60.31$$

you can conclude that the wire should be staked at 9 feet from the 12-foot pole.

Try It Exploration A

Video Video Video Video



You can confirm the minimum value of W with a graphing utility.
Figure 3.58

TECHNOLOGY From Example 4, you can see that applied optimization problems can involve a lot of algebra. If you have access to a graphing utility, you can confirm that $x = 9$ yields a minimum value of W by graphing

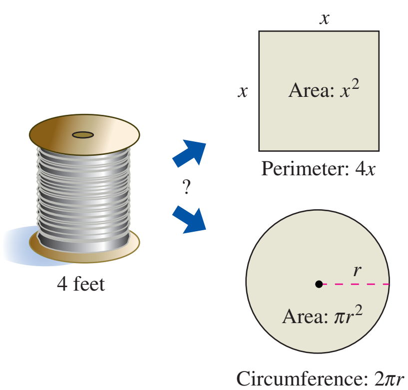
$$W = \sqrt{x^2 + 144} + \sqrt{x^2 - 60x + 1684}$$

as shown in Figure 3.58.

In each of the first four examples, the extreme value occurred at a critical number. Although this happens often, remember that an extreme value can also occur at an endpoint of an interval, as shown in Example 5.

EXAMPLE 5 An Endpoint Maximum

Four feet of wire is to be used to form a square and a circle. How much of the wire should be used for the square and how much should be used for the circle to enclose the maximum total area?



The quantity to be maximized is area: $A = x^2 + \pi r^2$.
Figure 3.59

Solution The total area (see Figure 3.59) is given by

$$A = (\text{area of square}) + (\text{area of circle})$$
$$A = x^2 + \pi r^2. \qquad \text{Primary equation}$$

Because the total length of wire is 4 feet, you obtain

$$4 = (\text{perimeter of square}) + (\text{circumference of circle})$$
$$4 = 4x + 2\pi r.$$

So, $r = 2(1 - x)/\pi$, and by substituting into the primary equation you have

$$\begin{aligned} A &= x^2 + \pi\left[\frac{2(1-x)}{\pi}\right]^2 \\ &= x^2 + \frac{4(1-x)^2}{\pi} \\ &= \frac{1}{\pi}[(\pi + 4)x^2 - 8x + 4]. \end{aligned}$$

The feasible domain is $0 \le x \le 1$ restricted by the square's perimeter. Because

$$\frac{dA}{dx} = \frac{2(\pi + 4)x - 8}{\pi}$$

the only critical number in $(0, 1)$ is $x = 4/(\pi + 4) \approx 0.56$. So, using

$$A(0) \approx 1.273, \qquad A(0.56) \approx 0.56, \qquad \text{and} \qquad A(1) = 1$$

you can conclude that the maximum area occurs when $x = 0$. That is, *all* the wire is used for the circle.

Try It | Exploration A | Exploration B | Exploration C

EXPLORATION

What would the answer be if Example 5 asked for the dimensions needed to enclose the *minimum* total area?

Let's review the primary equations developed in the first five examples. As applications go, these five examples are fairly simple, and yet the resulting primary equations are quite complicated.

$$V = 27x - \frac{x^3}{4} \qquad\qquad W = \sqrt{x^2 + 144} + \sqrt{x^2 - 60x + 1684}$$

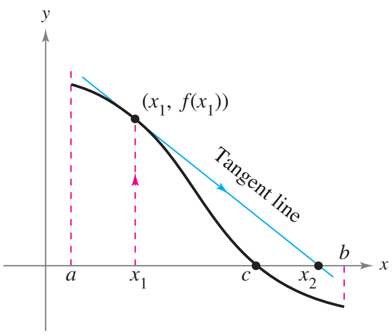
$$d = \sqrt{x^4 - 3x^2 + 4} \qquad\qquad A = \frac{1}{\pi}[(\pi + 4)x^2 - 8x + 4]$$

$$A = 30 + 2x + \frac{72}{x}$$

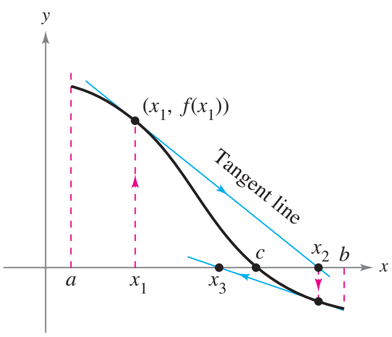
You must expect that real-life applications often involve equations that are *at least as complicated* as these five. Remember that one of the main goals of this course is to learn to use calculus to analyze equations that initially seem formidable.

Section 3.8 Newton's Method

- **Approximate a zero of a function using Newton's Method.**



(a)



(b)
The x-intercept of the tangent line approximates the zero of f.
Figure 3.60

Newton's Method

In this section you will study a technique for approximating the real zeros of a function. The technique is called **Newton's Method,** and it uses tangent lines to approximate the graph of the function near its x-intercepts.

To see how Newton's Method works, consider a function f that is continuous on the interval $[a, b]$ and differentiable on the interval (a, b). If $f(a)$ and $f(b)$ differ in sign, then, by the Intermediate Value Theorem, f must have at least one zero in the interval (a, b). Suppose you estimate this zero to occur at

$$x = x_1 \qquad \text{First estimate}$$

as shown in Figure 3.60(a). Newton's Method is based on the assumption that the graph of f and the tangent line at $(x_1, f(x_1))$ both cross the x-axis at *about* the same point. Because you can easily calculate the x-intercept for this tangent line, you can use it as a second (and, usually, better) estimate for the zero of f. The tangent line passes through the point $(x_1, f(x_1))$ with a slope of $f'(x_1)$. In point-slope form, the equation of the tangent line is therefore

$$y - f(x_1) = f'(x_1)(x - x_1)$$
$$y = f'(x_1)(x - x_1) + f(x_1).$$

Letting $y = 0$ and solving for x produces

$$x = x_1 - \frac{f(x_1)}{f'(x_1)}.$$

So, from the initial estimate x_1 you obtain a new estimate

$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)}. \qquad \text{Second estimate [see Figure 3.60(b)]}$$

You can improve on x_2 and calculate yet a third estimate

$$x_3 = x_2 - \frac{f(x_2)}{f'(x_2)}. \qquad \text{Third estimate}$$

Repeated application of this process is called Newton's Method.

Video

Newton's Method for Approximating the Zeros of a Function

Let $f(c) = 0$, where f is differentiable on an open interval containing c. Then, to approximate c, use the following steps.

1. Make an initial estimate x_1 that is close to c. (A graph is helpful.)
2. Determine a new approximation
$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.$$
3. If $|x_n - x_{n+1}|$ is within the desired accuracy, let x_{n+1} serve as the final approximation. Otherwise, return to Step 2 and calculate a new approximation.

Each successive application of this procedure is called an **iteration.**

Newton's Method

Isaac Newton first described the method for approximating the real zeros of a function in his text *Method of Fluxions*. Although the book was written in 1671, it was not published until 1736. Meanwhile, in 1690, Joseph Raphson (1648–1715) published a paper describing a method for approximating the real zeros of a function that was very similar to Newton's. For this reason, the method is often referred to as the Newton-Raphson method.

Technology

NOTE For many functions, just a few iterations of Newton's Method will produce approximations having very small errors, as shown in Example 1.

EXAMPLE 1 Using Newton's Method

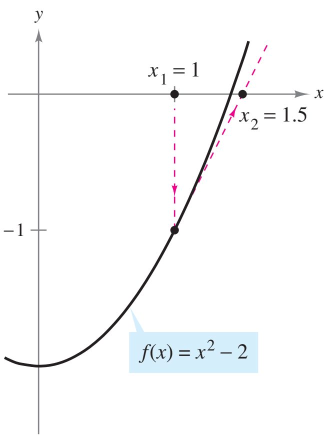
Calculate three iterations of Newton's Method to approximate a zero of $f(x) = x^2 - 2$. Use $x_1 = 1$ as the initial guess.

Solution Because $f(x) = x^2 - 2$, you have $f'(x) = 2x$, and the iterative process is given by the formula

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} = x_n - \frac{x_n^2 - 2}{2x_n}.$$

The calculations for three iterations are shown in the table.

n	x_n	$f(x_n)$	$f'(x_n)$	$\dfrac{f(x_n)}{f'(x_n)}$	$x_n - \dfrac{f(x_n)}{f'(x_n)}$
1	1.000000	−1.000000	2.000000	−0.500000	1.500000
2	1.500000	0.250000	3.000000	0.083333	1.416667
3	1.416667	0.006945	2.833334	0.002451	1.414216
4	1.414216				



The first iteration of Newton's Method
Figure 3.61

Editable Graph

Of course, in this case you know that the two zeros of the function are $\pm\sqrt{2}$. To six decimal places, $\sqrt{2} = 1.414214$. So, after only three iterations of Newton's Method, you have obtained an approximation that is within 0.000002 of an actual root. The first iteration of this process is shown in Figure 3.61.

Try It Exploration A Exploration B

EXAMPLE 2 Using Newton's Method

Use Newton's Method to approximate the zeros of

$$f(x) = 2x^3 + x^2 - x + 1.$$

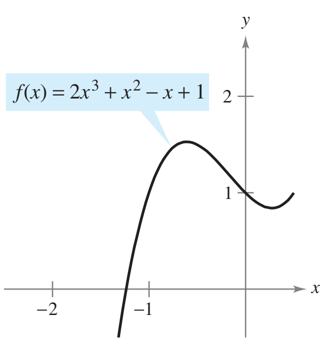
Continue the iterations until two successive approximations differ by less than 0.0001.

Solution Begin by sketching a graph of f, as shown in Figure 3.62. From the graph, you can observe that the function has only one zero, which occurs near $x = -1.2$. Next, differentiate f and form the iterative formula

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} = x_n - \frac{2x_n^3 + x_n^2 - x_n + 1}{6x_n^2 + 2x_n - 1}.$$

The calculations are shown in the table.

n	x_n	$f(x_n)$	$f'(x_n)$	$\dfrac{f(x_n)}{f'(x_n)}$	$x_n - \dfrac{f(x_n)}{f'(x_n)}$
1	−1.20000	0.18400	5.24000	0.03511	−1.23511
2	−1.23511	−0.00771	5.68276	−0.00136	−1.23375
3	−1.23375	0.00001	5.66533	0.00000	−1.23375
4	−1.23375				



After three iterations of Newton's Method, the zero of f is approximated to the desired accuracy.
Figure 3.62

Editable Graph

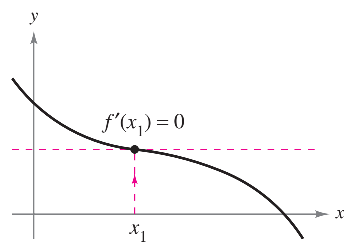
Because two successive approximations differ by less than the required 0.0001, you can estimate the zero of f to be -1.23375.

Try It Exploration A Exploration B

Open Exploration Video

When, as in Examples 1 and 2, the approximations approach a limit, the sequence $x_1, x_2, x_3, \ldots, x_n, \ldots$ is said to **converge.** Moreover, if the limit is c, it can be shown that c must be a zero of f.

Newton's Method does not always yield a convergent sequence. One way it can fail to do so is shown in Figure 3.63. Because Newton's Method involves division by $f'(x_n)$, it is clear that the method will fail if the derivative is zero for any x_n in the sequence. When you encounter this problem, you can usually overcome it by choosing a different value for x_1. Another way Newton's Method can fail is shown in the next example.



Newton's Method fails to converge if $f'(x_n) = 0$.
Figure 3.63

EXAMPLE 3 An Example in Which Newton's Method Fails

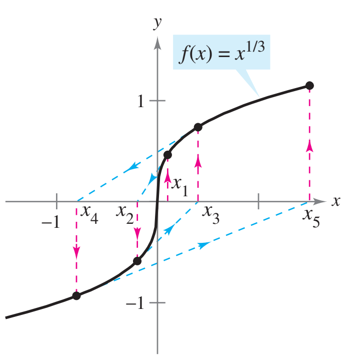
The function $f(x) = x^{1/3}$ is not differentiable at $x = 0$. Show that Newton's Method fails to converge using $x_1 = 0.1$.

Solution Because $f'(x) = \frac{1}{3}x^{-2/3}$, the iterative formula is

$$\begin{aligned} x_{n+1} &= x_n - \frac{f(x_n)}{f'(x_n)} \\ &= x_n - \frac{x_n^{1/3}}{\frac{1}{3}x_n^{-2/3}} \\ &= x_n - 3x_n \\ &= -2x_n. \end{aligned}$$

The calculations are shown in the table. This table and Figure 3.64 indicate that x_n continues to increase in magnitude as $n \to \infty$, and so the limit of the sequence does not exist.

n	x_n	$f(x_n)$	$f'(x_n)$	$\frac{f(x_n)}{f'(x_n)}$	$x_n - \frac{f(x_n)}{f'(x_n)}$
1	0.10000	0.46416	1.54720	0.30000	−0.20000
2	−0.20000	−0.58480	0.97467	−0.60000	0.40000
3	0.40000	0.73681	0.61401	1.20000	−0.80000
4	−0.80000	−0.92832	0.38680	−2.40000	1.60000



Newton's Method fails to converge for every x-value other than the actual zero of f.
Figure 3.64

Try It **Exploration A** **Exploration B**

NOTE In Example 3, the initial estimate $x_1 = 0.1$ fails to produce a convergent sequence. Try showing that Newton's Method also fails for every other choice of x_1 (other than the actual zero).

It can be shown that a condition sufficient to produce convergence of Newton's Method to a zero of f is that

$$\left|\frac{f(x)\,f''(x)}{[f'(x)]^2}\right| < 1$$

Condition for convergence

on an open interval containing the zero. For instance, in Example 1 this test would yield $f(x) = x^2 - 2, f'(x) = 2x, f''(x) = 2$, and

$$\left|\frac{f(x)\,f''(x)}{[f'(x)]^2}\right| = \left|\frac{(x^2-2)(2)}{4x^2}\right| = \left|\frac{1}{2} - \frac{1}{x^2}\right|.$$

Example 1

On the interval $(1, 3)$, this quantity is less than 1 and therefore the convergence of Newton's Method is guaranteed. On the other hand, in Example 3, you have $f(x) = x^{1/3}, f'(x) = \frac{1}{3}x^{-2/3}, f''(x) = -\frac{2}{9}x^{-5/3}$, and

$$\left|\frac{f(x)\,f''(x)}{[f'(x)]^2}\right| = \left|\frac{x^{1/3}(-2/9)(x^{-5/3})}{(1/9)(x^{-4/3})}\right| = 2$$

Example 3

which is not less than 1 for any value of x, so you cannot conclude that Newton's Method will converge.

Algebraic Solutions of Polynomial Equations

The zeros of some functions, such as

$$f(x) = x^3 - 2x^2 - x + 2$$

can be found by simple algebraic techniques, such as factoring. The zeros of other functions, such as

$$f(x) = x^3 - x + 1$$

cannot be found by *elementary* algebraic methods. This particular function has only one real zero, and by using more advanced algebraic techniques you can determine the zero to be

$$x = -\sqrt[3]{\frac{3 - \sqrt{23/3}}{6}} - \sqrt[3]{\frac{3 + \sqrt{23/3}}{6}}.$$

Because the *exact* solution is written in terms of square roots and cube roots, it is called a **solution by radicals.**

NOTE Try approximating the real zero of $f(x) = x^3 - x + 1$ and compare your result with the exact solution shown above.

NIELS HENRIK ABEL (1802–1829)

MathBio

EVARISTE GALOIS (1811–1832)

MathBio

Although the lives of both Abel and Galois were brief, their work in the fields of analysis and abstract algebra was far-reaching.

The determination of radical solutions of a polynomial equation is one of the fundamental problems of algebra. The earliest such result is the Quadratic Formula, which dates back at least to Babylonian times. The general formula for the zeros of a cubic function was developed much later. In the sixteenth century an Italian mathematician, Jerome Cardan, published a method for finding radical solutions to cubic and quartic equations. Then, for 300 years, the problem of finding a general quintic formula remained open. Finally, in the nineteenth century, the problem was answered independently by two young mathematicians. Niels Henrik Abel, a Norwegian mathematician, and Evariste Galois, a French mathematician, proved that it is not possible to solve a *general* fifth- (or higher-) degree polynomial equation by radicals. Of course, you can solve particular fifth-degree equations such as $x^5 - 1 = 0$, but Abel and Galois were able to show that no general *radical* solution exists.

Section 3.9 Differentials

- **Understand the concept of a tangent line approximation.**
- **Compare the value of the differential, dy, with the actual change in y, Δy.**
- **Estimate a propagated error using a differential.**
- **Find the differential of a function using differentiation formulas.**

EXPLORATION

Tangent Line Approximation Use a graphing utility to graph

$$f(x) = x^2.$$

In the same viewing window, graph the tangent line to the graph of f at the point $(1, 1)$. Zoom in twice on the point of tangency. Does your graphing utility distinguish between the two graphs? Use the *trace* feature to compare the two graphs. As the x-values get closer to 1, what can you say about the y-values?

Tangent Line Approximations

Newton's Method (Section 3.8) is an example of the use of a tangent line to a graph to approximate the graph. In this section, you will study other situations in which the graph of a function can be approximated by a straight line.

To begin, consider a function f that is differentiable at c. The equation for the tangent line at the point $(c, f(c))$ is given by

$$y - f(c) = f'(c)(x - c)$$
$$y = f(c) + f'(c)(x - c)$$

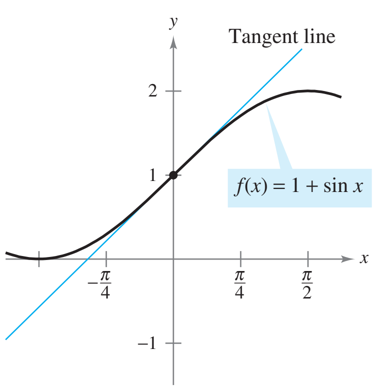
and is called the **tangent line approximation** (or **linear approximation**) **of f at c.** Because c is a constant, y is a linear function of x. Moreover, by restricting the values of x to be sufficiently close to c, the values of y can be used as approximations (to any desired accuracy) of the values of the function f. In other words, as $x \to c$, the limit of y is $f(c)$.

EXAMPLE 1 Using a Tangent Line Approximation

Find the tangent line approximation of

$$f(x) = 1 + \sin x$$

at the point $(0, 1)$. Then use a table to compare the y-values of the linear function with those of $f(x)$ on an open interval containing $x = 0$.



The tangent line approximation of f at the point $(0, 1)$
Figure 3.65

Editable Graph

Solution The derivative of f is

$$f'(x) = \cos x. \quad \text{First derivative}$$

So, the equation of the tangent line to the graph of f at the point $(0, 1)$ is

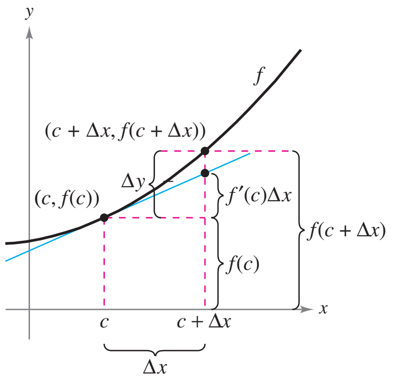
$$y - f(0) = f'(0)(x - 0)$$
$$y - 1 = (1)(x - 0)$$
$$y = 1 + x. \quad \text{Tangent line approximation}$$

The table compares the values of y given by this linear approximation with the values of $f(x)$ near $x = 0$. Notice that the closer x is to 0, the better the approximation is. This conclusion is reinforced by the graph shown in Figure 3.65.

x	-0.5	-0.1	-0.01	0	0.01	0.1	0.5
$f(x) = 1 + \sin x$	0.521	0.9002	0.9900002	1	1.0099998	1.0998	1.479
$y = 1 + x$	0.5	0.9	0.99	1	1.01	1.1	1.5

Try It | Exploration A | Open Exploration

NOTE Be sure you see that this linear approximation of $f(x) = 1 + \sin x$ depends on the point of tangency. At a different point on the graph of f, you would obtain a different tangent line approximation.



When Δx is small, $\Delta y = f(c + \Delta x) - f(c)$ is approximated by $f'(c)\Delta x$.
Figure 3.66

Differentials

When the tangent line to the graph of f at the point $(c, f(c))$

$$y = f(c) + f'(c)(x - c) \qquad \text{Tangent line at } (c, f(c))$$

is used as an approximation of the graph of f, the quantity $x - c$ is called the change in x, and is denoted by Δx, as shown in Figure 3.66. When Δx is small, the change in y (denoted by Δy) can be approximated as shown.

$$\Delta y = f(c + \Delta x) - f(c) \qquad \text{Actual change in } y$$
$$\approx f'(c)\Delta x \qquad \text{Approximate change in } y$$

For such an approximation, the quantity Δx is traditionally denoted by dx, and is called the **differential of *x*.** The expression $f'(x)\,dx$ is denoted by dy, and is called the **differential of *y*.**

Definition of Differentials

Let $y = f(x)$ represent a function that is differentiable on an open interval containing x. The **differential of *x*** (denoted by dx) is any nonzero real number. The **differential of *y*** (denoted by dy) is

$$dy = f'(x)\,dx.$$

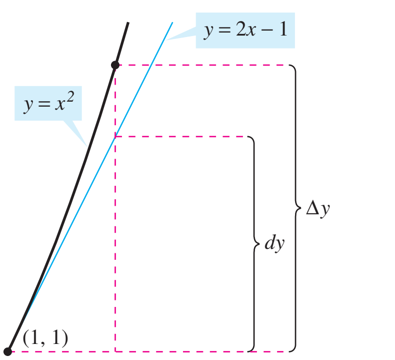
Video

In many types of applications, the differential of y can be used as an approximation of the change in y. That is,

$$\Delta y \approx dy \qquad \text{or} \qquad \Delta y \approx f'(x)dx.$$

EXAMPLE 2 Comparing Δy and dy

Let $y = x^2$. Find dy when $x = 1$ and $dx = 0.01$. Compare this value with Δy for $x = 1$ and $\Delta x = 0.01$.



The change in y, Δy, is approximated by the differential of y, dy.
Figure 3.67

Solution Because $y = f(x) = x^2$, you have $f'(x) = 2x$, and the differential dy is given by

$$dy = f'(x)\,dx = f'(1)(0.01) = 2(0.01) = 0.02. \qquad \text{Differential of } y$$

Now, using $\Delta x = 0.01$, the change in y is

$$\Delta y = f(x + \Delta x) - f(x) = f(1.01) - f(1) = (1.01)^2 - 1^2 = 0.0201.$$

Figure 3.67 shows the geometric comparison of dy and Δy. Try comparing other values of dy and Δy. You will see that the values become closer to each other as dx (or Δx) approaches 0.

Try It Exploration A

In Example 2, the tangent line to the graph of $f(x) = x^2$ at $x = 1$ is

$$y = 2x - 1 \qquad \text{or} \qquad g(x) = 2x - 1. \qquad \text{Tangent line to the graph of } f \text{ at } x = 1.$$

For x-values near 1, this line is close to the graph of f, as shown in Figure 3.67. For instance,

$$f(1.01) = 1.01^2 = 1.0201 \qquad \text{and} \qquad g(1.01) = 2(1.01) - 1 = 1.02.$$

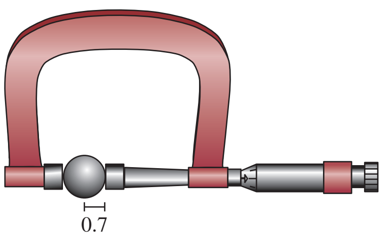
Error Propagation

Physicists and engineers tend to make liberal use of the approximation of Δy by dy. One way this occurs in practice is in the estimation of errors propagated by physical measuring devices. For example, if you let x represent the measured value of a variable and let $x + \Delta x$ represent the exact value, then Δx is the *error in measurement*. Finally, if the measured value x is used to compute another value $f(x)$, the difference between $f(x + \Delta x)$ and $f(x)$ is the **propagated error.**

$$\underbrace{f(x + \overbrace{\Delta x}^{\text{Measurement error}})}_{\text{Exact value}} - \underbrace{f(x)}_{\text{Measured value}} = \overbrace{\Delta y}^{\text{Propagated error}}$$

EXAMPLE 3 Estimation of Error

The radius of a ball bearing is measured to be 0.7 inch, as shown in Figure 3.68. If the measurement is correct to within 0.01 inch, estimate the propagated error in the volume V of the ball bearing.



Ball bearing with measured radius that is correct to within 0.01 inch
Figure 3.68

Solution The formula for the volume of a sphere is $V = \frac{4}{3}\pi r^3$, where r is the radius of the sphere. So, you can write

$$r = 0.7 \qquad \text{Measured radius}$$

and

$$-0.01 \le \Delta r \le 0.01. \qquad \text{Possible error}$$

To approximate the propagated error in the volume, differentiate V to obtain $dV/dr = 4\pi r^2$ and write

$$\begin{aligned} \Delta V &\approx dV && \text{Approximate } \Delta V \text{ by } dV. \\ &= 4\pi r^2\, dr \\ &= 4\pi(0.7)^2(\pm 0.01) && \text{Substitute for } r \text{ and } dr. \\ &\approx \pm 0.06158 \text{ cubic inch.} \end{aligned}$$

So, the volume has a propagated error of about 0.06 cubic inch.

Try It | Exploration A | Exploration B

Video | Video

Would you say that the propagated error in Example 3 is large or small? The answer is best given in *relative* terms by comparing dV with V. The ratio

$$\begin{aligned} \frac{dV}{V} &= \frac{4\pi r^2\, dr}{\frac{4}{3}\pi r^3} && \text{Ratio of } dV \text{ to } V \\ &= \frac{3\, dr}{r} && \text{Simplify.} \\ &\approx \frac{3}{0.7}(\pm 0.01) && \text{Substitute for } dr \text{ and } r. \\ &\approx \pm 0.0429 \end{aligned}$$

is called the **relative error.** The corresponding **percent error** is approximately 4.29%.

Calculating Differentials

Each of the differentiation rules that you studied in Chapter 2 can be written in **differential form.** For example, suppose u and v are differentiable functions of x. By the definition of differentials, you have

$$du = u'\,dx \quad \text{and} \quad dv = v'\,dx.$$

So, you can write the differential form of the Product Rule as shown below.

$$\begin{aligned} d[uv] &= \frac{d}{dx}[uv]\,dx && \text{Differential of } uv \\ &= [uv' + vu']\,dx && \text{Product Rule} \\ &= uv'\,dx + vu'\,dx \\ &= u\,dv + v\,du \end{aligned}$$

Differential Formulas

Let u and v be differentiable functions of x.

Constant multiple: $d[cu] = c\,du$

Sum or difference: $d[u \pm v] = du \pm dv$

Product: $d[uv] = u\,dv + v\,du$

Quotient: $d\left[\dfrac{u}{v}\right] = \dfrac{v\,du - u\,dv}{v^2}$

EXAMPLE 4 Finding Differentials

	Function	*Derivative*	*Differential*
a.	$y = x^2$	$\dfrac{dy}{dx} = 2x$	$dy = 2x\,dx$
b.	$y = 2\sin x$	$\dfrac{dy}{dx} = 2\cos x$	$dy = 2\cos x\,dx$
c.	$y = x\cos x$	$\dfrac{dy}{dx} = -x\sin x + \cos x$	$dy = (-x\sin x + \cos x)\,dx$
d.	$y = \dfrac{1}{x}$	$\dfrac{dy}{dx} = -\dfrac{1}{x^2}$	$dy = -\dfrac{dx}{x^2}$

Try It Exploration A

The notation in Example 4 is called the **Leibniz notation** for derivatives and differentials, named after the German mathematician Gottfried Wilhelm Leibniz. The beauty of this notation is that it provides an easy way to remember several important calculus formulas by making it seem as though the formulas were derived from algebraic manipulations of differentials. For instance, in Leibniz notation, the *Chain Rule*

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

would appear to be true because the du's divide out. Even though this reasoning is *incorrect*, the notation does help one remember the Chain Rule.

Gottfried Wilhelm Leibniz (1646–1716)

Both Leibniz and Newton are credited with creating calculus. It was Leibniz, however, who tried to broaden calculus by developing rules and formal notation. He often spent days choosing an appropriate notation for a new concept.

MathBio

EXAMPLE 5 Finding the Differential of a Composite Function

$$y = f(x) = \sin 3x \qquad \text{Original function}$$
$$f'(x) = 3\cos 3x \qquad \text{Apply Chain Rule.}$$
$$dy = f'(x)\,dx = 3\cos 3x\,dx \qquad \text{Differential form}$$

Try It Exploration A

EXAMPLE 6 Finding the Differential of a Composite Function

$$y = f(x) = (x^2 + 1)^{1/2} \qquad \text{Original function}$$
$$f'(x) = \frac{1}{2}(x^2 + 1)^{-1/2}(2x) = \frac{x}{\sqrt{x^2 + 1}} \qquad \text{Apply Chain Rule.}$$
$$dy = f'(x)\,dx = \frac{x}{\sqrt{x^2 + 1}}\,dx \qquad \text{Differential form}$$

Try It Exploration A Exploration B Technology

Differentials can be used to approximate function values. To do this for the function given by $y = f(x)$, you use the formula

$$f(x + \Delta x) \approx f(x) + dy = f(x) + f'(x)\,dx$$

which is derived from the approximation $\Delta y = f(x + \Delta x) - f(x) \approx dy$. The key to using this formula is to choose a value for x that makes the calculations easier, as shown in Example 7.

EXAMPLE 7 Approximating Function Values

Use differentials to approximate $\sqrt{16.5}$.

Solution Using $f(x) = \sqrt{x}$, you can write

$$f(x + \Delta x) \approx f(x) + f'(x)\,dx = \sqrt{x} + \frac{1}{2\sqrt{x}}\,dx.$$

Now, choosing $x = 16$ and $dx = 0.5$, you obtain the following approximation.

$$f(x + \Delta x) = \sqrt{16.5} \approx \sqrt{16} + \frac{1}{2\sqrt{16}}(0.5) = 4 + \left(\frac{1}{8}\right)\left(\frac{1}{2}\right) = 4.0625$$

Try It Exploration A

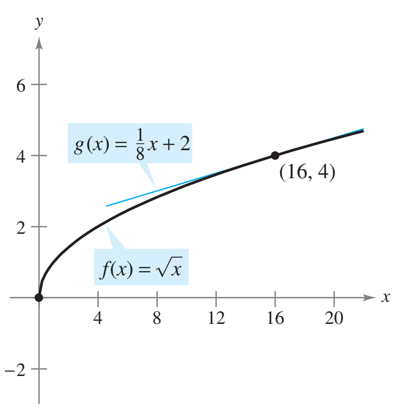


Figure 3.69

The tangent line approximation to $f(x) = \sqrt{x}$ at $x = 16$ is the line $g(x) = \frac{1}{8}x + 2$. For x-values near 16, the graphs of f and g are close together, as shown in Figure 3.69. For instance,

$$f(16.5) = \sqrt{16.5} \approx 4.0620 \quad \text{and} \quad g(16.5) = \frac{1}{8}(16.5) + 2 = 4.0625.$$

In fact, if you use a graphing utility to zoom in near the point of tangency (16, 4), you will see that the two graphs appear to coincide. Notice also that as you move farther away from the point of tangency, the linear approximation is less accurate.

Section 4.1 Antiderivatives and Indefinite Integration

- **Write the general solution of a differential equation.**
- **Use indefinite integral notation for antiderivatives.**
- **Use basic integration rules to find antiderivatives.**
- **Find a particular solution of a differential equation.**

EXPLORATION

Finding Antiderivatives For each derivative, describe the original function F.

a. $F'(x) = 2x$

b. $F'(x) = x$

c. $F'(x) = x^2$

d. $F'(x) = \dfrac{1}{x^2}$

e. $F'(x) = \dfrac{1}{x^3}$

f. $F'(x) = \cos x$

What strategy did you use to find F?

Antiderivatives

Suppose you were asked to find a function F whose derivative is $f(x) = 3x^2$. From your knowledge of derivatives, you would probably say that

$$F(x) = x^3 \text{ because } \frac{d}{dx}[x^3] = 3x^2.$$

The function F is an *antiderivative* of f.

Definition of an Antiderivative

A function F is an **antiderivative** of f on an interval I if $F'(x) = f(x)$ for all x in I.

Video

Note that F is called *an* antiderivative of f, rather than *the* antiderivative of f. To see why, observe that

$$F_1(x) = x^3, \quad F_2(x) = x^3 - 5, \quad \text{and} \quad F_3(x) = x^3 + 97$$

are all antiderivatives of $f(x) = 3x^2$. In fact, for any constant C, the function given by $F(x) = x^3 + C$ is an antiderivative of f.

THEOREM 4.1 Representation of Antiderivatives

If F is an antiderivative of f on an interval I, then G is an antiderivative of f on the interval I if and only if G is of the form $G(x) = F(x) + C$, for all x in I where C is a constant.

Proof The proof of Theorem 4.1 in one direction is straightforward. That is, if $G(x) = F(x) + C$, $F'(x) = f(x)$, and C is a constant, then

$$G'(x) = \frac{d}{dx}[F(x) + C] = F'(x) + 0 = f(x).$$

To prove this theorem in the other direction, assume that G is an antiderivative of f. Define a function H such that

$$H(x) = G(x) - F(x).$$

If H is not constant on the interval I, there must exist a and b $(a < b)$ in the interval such that $H(a) \neq H(b)$. Moreover, because H is differentiable on (a, b), you can apply the Mean Value Theorem to conclude that there exists some c in (a, b) such that

$$H'(c) = \frac{H(b) - H(a)}{b - a}.$$

Because $H(b) \neq H(a)$, it follows that $H'(c) \neq 0$. However, because $G'(c) = F'(c)$, you know that $H'(c) = G'(c) - F'(c) = 0$, which contradicts the fact that $H'(c) \neq 0$. Consequently, you can conclude that $H(x)$ is a constant, C. So, $G(x) - F(x) = C$ and it follows that $G(x) = F(x) + C$.

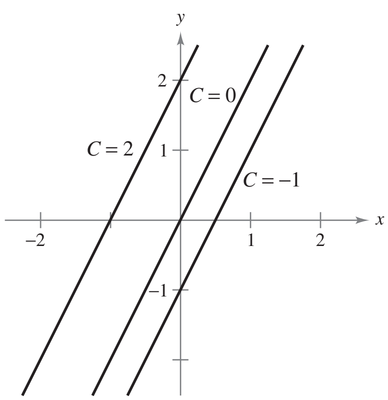
Using Theorem 4.1, you can represent the entire family of antiderivatives of a function by adding a constant to a *known* antiderivative. For example, knowing that $D_x[x^2] = 2x$, you can represent the family of *all* antiderivatives of $f(x) = 2x$ by

$$G(x) = x^2 + C \qquad \text{Family of all antiderivatives of } f(x) = 2x$$

where C is a constant. The constant C is called the **constant of integration.** The family of functions represented by G is the **general antiderivative** of f, and $G(x) = x^2 + C$ is the **general solution** of the *differential equation*

$$G'(x) = 2x. \qquad \text{Differential equation}$$

A **differential equation** in x and y is an equation that involves x, y, and derivatives of y. For instance, $y' = 3x$ and $y' = x^2 + 1$ are examples of differential equations.



Functions of the form $y = 2x + C$
Figure 4.1

Editable Graph

Video

EXAMPLE 1 Solving a Differential Equation

Find the general solution of the differential equation $y' = 2$.

Solution To begin, you need to find a function whose derivative is 2. One such function is

$$y = 2x. \qquad 2x \text{ is } \textit{an} \text{ antiderivative of } 2.$$

Now, you can use Theorem 4.1 to conclude that the general solution of the differential equation is

$$y = 2x + C. \qquad \text{General solution}$$

The graphs of several functions of the form $y = 2x + C$ are shown in Figure 4.1.

Try It

Exploration A

Notation for Antiderivatives

When solving a differential equation of the form

$$\frac{dy}{dx} = f(x)$$

it is convenient to write it in the equivalent differential form

$$dy = f(x)\,dx.$$

The operation of finding all solutions of this equation is called **antidifferentiation** (or **indefinite integration**) and is denoted by an integral sign $\int$. The general solution is denoted by

Variable of integration

Constant of integration

$$y = \int f(x)\,dx = F(x) + C.$$

Integrand

NOTE In this text, the notation $\int f(x)\,dx = F(x) + C$ means that F is an antiderivative of f *on an interval*.

The expression $\int f(x)\,dx$ is read as the *antiderivative of f with respect to x*. So, the differential dx serves to identify x as the variable of integration. The term **indefinite integral** is a synonym for antiderivative.

Basic Integration Rules

The inverse nature of integration and differentiation can be verified by substituting $F'(x)$ for $f(x)$ in the indefinite integration definition to obtain

$$\int F'(x)\,dx = F(x) + C.$$

Integration is the "inverse" of differentiation.

Moreover, if $\int f(x)\,dx = F(x) + C$, then

$$\frac{d}{dx}\left[\int f(x)\,dx\right] = f(x).$$

Differentiation is the "inverse" of integration.

These two equations allow you to obtain integration formulas directly from differentiation formulas, as shown in the following summary.

Basic Integration Rules

Differentiation Formula	*Integration Formula*
$\frac{d}{dx}[C] = 0$	$\int 0\,dx = C$
$\frac{d}{dx}[kx] = k$	$\int k\,dx = kx + C$
$\frac{d}{dx}[kf(x)] = kf'(x)$	$\int kf(x)\,dx = k\int f(x)\,dx$
$\frac{d}{dx}[f(x) \pm g(x)] = f'(x) \pm g'(x)$	$\int [f(x) \pm g(x)]\,dx = \int f(x)\,dx \pm \int g(x)\,dx$
$\frac{d}{dx}[x^n] = nx^{n-1}$	$\int x^n\,dx = \frac{x^{n+1}}{n+1} + C, \quad n \neq -1$ Power Rule
$\frac{d}{dx}[\sin x] = \cos x$	$\int \cos x\,dx = \sin x + C$
$\frac{d}{dx}[\cos x] = -\sin x$	$\int \sin x\,dx = -\cos x + C$
$\frac{d}{dx}[\tan x] = \sec^2 x$	$\int \sec^2 x\,dx = \tan x + C$
$\frac{d}{dx}[\sec x] = \sec x \tan x$	$\int \sec x \tan x\,dx = \sec x + C$
$\frac{d}{dx}[\cot x] = -\csc^2 x$	$\int \csc^2 x\,dx = -\cot x + C$
$\frac{d}{dx}[\csc x] = -\csc x \cot x$	$\int \csc x \cot x\,dx = -\csc x + C$

NOTE Note that the Power Rule for Integration has the restriction that $n \neq -1$. The evaluation of $\int 1/x\,dx$ must wait until the introduction of the natural logarithm function in Chapter 5.

EXAMPLE 2 Applying the Basic Integration Rules

Describe the antiderivatives of $3x$.

Solution

$$\begin{aligned}\int 3x\,dx &= 3\int x\,dx && \text{Constant Multiple Rule}\\ &= 3\int x^1\,dx && \text{Rewrite } x \text{ as } x^1.\\ &= 3\left(\frac{x^2}{2}\right) + C && \text{Power Rule } (n = 1)\\ &= \frac{3}{2}x^2 + C && \text{Simplify.}\end{aligned}$$

So, the antiderivatives of $3x$ are of the form $\frac{3}{2}x^2 + C$, where C is any constant.

Try It | Exploration A

When indefinite integrals are evaluated, a strict application of the basic integration rules tends to produce complicated constants of integration. For instance, in Example 2, you could have written

$$\int 3x\,dx = 3\int x\,dx = 3\left(\frac{x^2}{2} + C\right) = \frac{3}{2}x^2 + 3C.$$

However, because C represents *any* constant, it is both cumbersome and unnecessary to write $3C$ as the constant of integration. So, $\frac{3}{2}x^2 + 3C$ is written in the simpler form, $\frac{3}{2}x^2 + C$.

In Example 2, note that the general pattern of integration is similar to that of differentiation.

Original integral ⇨ Rewrite ⇨ Integrate ⇨ Simplify

TECHNOLOGY Some software programs, such as *Derive*, *Maple*, *Mathcad*, *Mathematica*, and the *TI-89*, are capable of performing integration symbolically. If you have access to such a symbolic integration utility, try using it to evaluate the indefinite integrals in Example 3.

EXAMPLE 3 Rewriting Before Integrating

	Original Integral	*Rewrite*	*Integrate*	*Simplify*
a.	$\int \frac{1}{x^3}\,dx$	$\int x^{-3}\,dx$	$\frac{x^{-2}}{-2} + C$	$-\frac{1}{2x^2} + C$
b.	$\int \sqrt{x}\,dx$	$\int x^{1/2}\,dx$	$\frac{x^{3/2}}{3/2} + C$	$\frac{2}{3}x^{3/2} + C$
c.	$\int 2\sin x\,dx$	$2\int \sin x\,dx$	$2(-\cos x) + C$	$-2\cos x + C$

Try It | Exploration A | Open Exploration | Video

Remember that you can check your answer to an antidifferentiation problem by differentiating. For instance, in Example 3(b), you can check that $\frac{2}{3}x^{3/2} + C$ is the correct antiderivative by differentiating the answer to obtain

$$D_x\left[\frac{2}{3}x^{3/2} + C\right] = \left(\frac{2}{3}\right)\left(\frac{3}{2}\right)x^{1/2} = \sqrt{x}. \qquad \text{Use differentiation to check antiderivative.}$$

The basic integration rules listed earlier in this section allow you to integrate any polynomial function, as shown in Example 4.

EXAMPLE 4 Integrating Polynomial Functions

a. $\int dx = \int 1\,dx$ — Integrand is understood to be 1.

$= x + C$ — Integrate.

b. $\int (x + 2)\,dx = \int x\,dx + \int 2\,dx$

$= \dfrac{x^2}{2} + C_1 + 2x + C_2$ — Integrate.

$= \dfrac{x^2}{2} + 2x + C$ — $C = C_1 + C_2$

The second line in the solution is usually omitted.

c. $\int (3x^4 - 5x^2 + x)\,dx = 3\left(\dfrac{x^5}{5}\right) - 5\left(\dfrac{x^3}{3}\right) + \dfrac{x^2}{2} + C$ — Integrate.

$= \dfrac{3}{5}x^5 - \dfrac{5}{3}x^3 + \dfrac{1}{2}x^2 + C$ — Simplify.

Try It   Exploration A

EXAMPLE 5 Rewriting Before Integrating

$\int \dfrac{x + 1}{\sqrt{x}}\,dx = \int \left(\dfrac{x}{\sqrt{x}} + \dfrac{1}{\sqrt{x}}\right) dx$ — Rewrite as two fractions.

$= \int (x^{1/2} + x^{-1/2})\,dx$ — Rewrite with fractional exponents.

$= \dfrac{x^{3/2}}{3/2} + \dfrac{x^{1/2}}{1/2} + C$ — Integrate.

$= \dfrac{2}{3}x^{3/2} + 2x^{1/2} + C$ — Simplify.

$= \dfrac{2}{3}\sqrt{x}(x + 3) + C$

Try It   Exploration A   Exploration B

NOTE When integrating quotients, do not integrate the numerator and denominator separately. This is no more valid in integration than it is in differentiation. For instance, in Example 5, be sure you understand that

$$\int \frac{x + 1}{\sqrt{x}}\,dx = \frac{2}{3}\sqrt{x}(x + 3) + C \text{ is not the same as } \frac{\int (x + 1)\,dx}{\int \sqrt{x}\,dx} = \frac{\frac{1}{2}x^2 + x + C_1}{\frac{2}{3}x\sqrt{x} + C_2}.$$

EXAMPLE 6 Rewriting Before Integrating

$\int \dfrac{\sin x}{\cos^2 x}\,dx = \int \left(\dfrac{1}{\cos x}\right)\left(\dfrac{\sin x}{\cos x}\right) dx$ — Rewrite as a product.

$= \int \sec x \tan x\,dx$ — Rewrite using trigonometric identities.

$= \sec x + C$ — Integrate.

Try It   Exploration A

Initial Conditions and Particular Solutions

You have already seen that the equation $y = \int f(x)\,dx$ has many solutions (each differing from the others by a constant). This means that the graphs of any two antiderivatives of f are vertical translations of each other. For example, Figure 4.2 shows the graphs of several antiderivatives of the form

$$y = \int (3x^2 - 1)\,dx = x^3 - x + C \qquad \text{General solution}$$

for various integer values of C. Each of these antiderivatives is a solution of the differential equation

$$\frac{dy}{dx} = 3x^2 - 1.$$

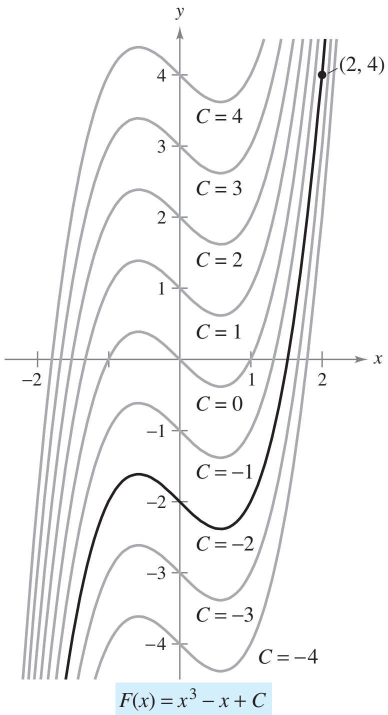
In many applications of integration, you are given enough information to determine a **particular solution.** To do this, you need only know the value of $y = F(x)$ for one value of x. This information is called an **initial condition.** For example, in Figure 4.2, only one curve passes through the point $(2, 4)$. To find this curve, you can use the following information.

$$F(x) = x^3 - x + C \qquad \text{General solution}$$

$$F(2) = 4 \qquad \text{Initial condition}$$

By using the initial condition in the general solution, you can determine that $F(2) = 8 - 2 + C = 4$, which implies that $C = -2$. So, you obtain

$$F(x) = x^3 - x - 2. \qquad \text{Particular solution}$$



The particular solution that satisfies the initial condition $F(2) = 4$ is $F(x) = x^3 - x - 2$.
Figure 4.2

Animation

EXAMPLE 7 Finding a Particular Solution

Find the general solution of

$$F'(x) = \frac{1}{x^2}, \quad x > 0$$

and find the particular solution that satisfies the initial condition $F(1) = 0$.

Solution To find the general solution, integrate to obtain

$$F(x) = \int \frac{1}{x^2}\,dx \qquad F(x) = \int F'(x)dx$$

$$= \int x^{-2}\,dx \qquad \text{Rewrite as a power.}$$

$$= \frac{x^{-1}}{-1} + C \qquad \text{Integrate.}$$

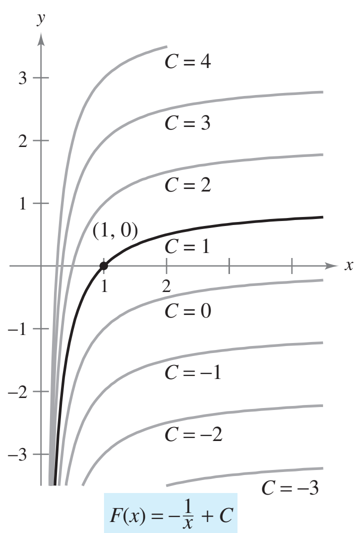
$$= -\frac{1}{x} + C, \quad x > 0. \qquad \text{General solution}$$

Using the initial condition $F(1) = 0$, you can solve for C as follows.

$$F(1) = -\frac{1}{1} + C = 0 \quad \Rightarrow \quad C = 1$$

So, the particular solution, as shown in Figure 4.3, is

$$F(x) = -\frac{1}{x} + 1, \quad x > 0. \qquad \text{Particular solution}$$



The particular solution that satisfies the initial condition $F(1) = 0$ is $F(x) = -(1/x) + 1, x > 0$.
Figure 4.3

Editable Graph

Try It

Exploration A

So far in this section you have been using x as the variable of integration. In applications, it is often convenient to use a different variable. For instance, in the following example involving *time*, the variable of integration is t.

EXAMPLE 8 Solving a Vertical Motion Problem

A ball is thrown upward with an initial velocity of 64 feet per second from an initial height of 80 feet.

a. Find the position function giving the height s as a function of the time t.

b. When does the ball hit the ground?

Solution

a. Let $t = 0$ represent the initial time. The two given initial conditions can be written as follows.

$$s(0) = 80 \qquad \text{Initial height is 80 feet.}$$
$$s'(0) = 64 \qquad \text{Initial velocity is 64 feet per second.}$$

Using -32 feet per second per second as the acceleration due to gravity, you can write

$$s''(t) = -32$$
$$s'(t) = \int s''(t)\,dt = \int -32\,dt = -32t + C_1.$$

Using the initial velocity, you obtain $s'(0) = 64 = -32(0) + C_1$, which implies that $C_1 = 64$. Next, by integrating $s'(t)$, you obtain

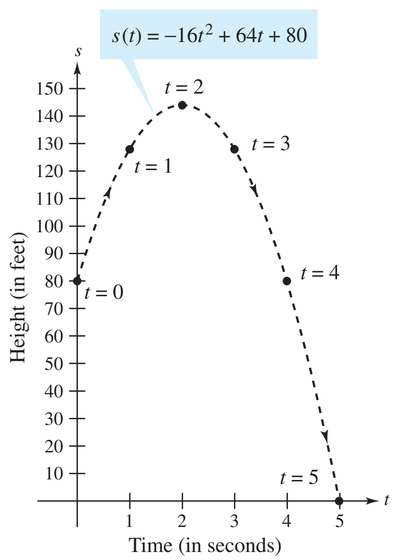
$$s(t) = \int s'(t)\,dt = \int (-32t + 64)\,dt = -16t^2 + 64t + C_2.$$

Using the initial height, you obtain

$$s(0) = 80 = -16(0^2) + 64(0) + C_2$$

which implies that $C_2 = 80$. So, the position function is

$$s(t) = -16t^2 + 64t + 80. \qquad \text{See Figure 4.4.}$$



Height of a ball at time t
Figure 4.4

Animation

b. Using the position function found in part (a), you can find the time that the ball hits the ground by solving the equation $s(t) = 0$.

$$s(t) = -16t^2 + 64t + 80 = 0$$
$$-16(t + 1)(t - 5) = 0$$
$$t = -1, 5$$

Because t must be positive, you can conclude that the ball hits the ground 5 seconds after it was thrown.

NOTE In Example 8, note that the position function has the form

$$s(t) = \tfrac{1}{2}gt^2 + v_0t + s_0$$

where $g = -32$, v_0 is the initial velocity, and s_0 is the initial height, as presented in Section 2.2.

Try It | Exploration A

Example 8 shows how to use calculus to analyze vertical motion problems in which the acceleration is determined by a gravitational force. You can use a similar strategy to analyze other linear motion problems (vertical or horizontal) in which the acceleration (or deceleration) is the result of some other force, as you will see in Exercises 77–86.

Before you begin the exercise set, be sure you realize that one of the most important steps in integration is *rewriting the integrand* in a form that fits the basic integration rules. To illustrate this point further, here are some additional examples.

Original Integral	*Rewrite*	*Integrate*	*Simplify*
$\displaystyle\int \frac{2}{\sqrt{x}}\,dx$	$\displaystyle 2\int x^{-1/2}\,dx$	$\displaystyle 2\left(\frac{x^{1/2}}{1/2}\right) + C$	$4x^{1/2} + C$
$\displaystyle\int (t^2 + 1)^2\,dt$	$\displaystyle\int (t^4 + 2t^2 + 1)\,dt$	$\displaystyle \frac{t^5}{5} + 2\left(\frac{t^3}{3}\right) + t + C$	$\displaystyle \frac{1}{5}t^5 + \frac{2}{3}t^3 + t + C$
$\displaystyle\int \frac{x^3 + 3}{x^2}\,dx$	$\displaystyle\int (x + 3x^{-2})\,dx$	$\displaystyle \frac{x^2}{2} + 3\left(\frac{x^{-1}}{-1}\right) + C$	$\displaystyle \frac{1}{2}x^2 - \frac{3}{x} + C$
$\displaystyle\int \sqrt[3]{x}(x - 4)\,dx$	$\displaystyle\int (x^{4/3} - 4x^{1/3})\,dx$	$\displaystyle \frac{x^{7/3}}{7/3} - 4\left(\frac{x^{4/3}}{4/3}\right) + C$	$\displaystyle \frac{3}{7}x^{7/3} - 3x^{4/3}$

Section 4.2 Area

- **Use sigma notation to write and evaluate a sum.**
- **Understand the concept of area.**
- **Approximate the area of a plane region.**
- **Find the area of a plane region using limits.**

Sigma Notation

In the preceding section, you studied antidifferentiation. In this section, you will look further into a problem introduced in Section 1.1—that of finding the area of a region in the plane. At first glance, these two ideas may seem unrelated, but you will discover in Section 4.4 that they are closely related by an extremely important theorem called the Fundamental Theorem of Calculus.

This section begins by introducing a concise notation for sums. This notation is called **sigma notation** because it uses the uppercase Greek letter sigma, written as Σ.

> **Sigma Notation**
>
> The sum of n terms $a_1, a_2, a_3, \ldots, a_n$ is written as
>
> $$\sum_{i=1}^{n} a_i = a_1 + a_2 + a_3 + \cdots + a_n$$
>
> where i is the **index of summation,** a_i is the ***i*th term** of the sum, and the **upper and lower bounds of summation** are n and 1.

NOTE The upper and lower bounds must be constant with respect to the index of summation. However, the lower bound doesn't have to be 1. Any integer less than or equal to the upper bound is legitimate.

EXAMPLE 1 Examples of Sigma Notation

a. $\displaystyle\sum_{i=1}^{6} i = 1 + 2 + 3 + 4 + 5 + 6$

b. $\displaystyle\sum_{i=0}^{5} (i + 1) = 1 + 2 + 3 + 4 + 5 + 6$

c. $\displaystyle\sum_{j=3}^{7} j^2 = 3^2 + 4^2 + 5^2 + 6^2 + 7^2$

d. $\displaystyle\sum_{k=1}^{n} \frac{1}{n}(k^2 + 1) = \frac{1}{n}(1^2 + 1) + \frac{1}{n}(2^2 + 1) + \cdots + \frac{1}{n}(n^2 + 1)$

e. $\displaystyle\sum_{i=1}^{n} f(x_i)\,\Delta x = f(x_1)\,\Delta x + f(x_2)\,\Delta x + \cdots + f(x_n)\,\Delta x$

From parts (a) and (b), notice that the same sum can be represented in different ways using sigma notation.

Try It | Exploration A | Technology

FOR FURTHER INFORMATION For a geometric interpretation of summation formulas, see the article, "Looking at $\sum_{k=1}^{n} k$ and $\sum_{k=1}^{n} k^2$ Geometrically" by Eric Hegblom in *Mathematics Teacher*.

MathArticle

Although any variable can be used as the index of summation i, j, and k are often used. Notice in Example 1 that the index of summation does not appear in the terms of the expanded sum.

THE SUM OF THE FIRST 100 INTEGERS

Carl Friedrich Gauss's (1777–1855) teacher asked him to add all the integers from 1 to 100. When Gauss returned with the correct answer after only a few moments, the teacher could only look at him in astounded silence. This is what Gauss did:

$$\begin{array}{rcrcrcccr} 1 & + & 2 & + & 3 & + & \cdots & + & 100 \\ 100 & + & 99 & + & 98 & + & \cdots & + & 1 \\ \hline 101 & + & 101 & + & 101 & + & \cdots & + & 101 \end{array}$$

$$\frac{100 \times 101}{2} = 5050$$

This is generalized by Theorem 4.2, where

$$\sum_{i=1}^{100} i = \frac{100(101)}{2} = 5050.$$

The following properties of summation can be derived using the associative and commutative properties of addition and the distributive property of addition over multiplication. (In the first property, k is a constant.)

1. $\displaystyle\sum_{i=1}^{n} ka_i = k\sum_{i=1}^{n} a_i$
2. $\displaystyle\sum_{i=1}^{n} (a_i \pm b_i) = \sum_{i=1}^{n} a_i \pm \sum_{i=1}^{n} b_i$

The next theorem lists some useful formulas for sums of powers. A proof of this theorem is given in Appendix A.

THEOREM 4.2 Summation Formulas

1. $\displaystyle\sum_{i=1}^{n} c = cn$
2. $\displaystyle\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$
3. $\displaystyle\sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}$
4. $\displaystyle\sum_{i=1}^{n} i^3 = \frac{n^2(n+1)^2}{4}$

EXAMPLE 2 Evaluating a Sum

Evaluate $\displaystyle\sum_{i=1}^{n} \frac{i+1}{n^2}$ for $n = 10, 100, 1000$, and $10{,}000$.

Solution Applying Theorem 4.2, you can write

$$\begin{aligned} \sum_{i=1}^{n} \frac{i+1}{n^2} &= \frac{1}{n^2}\sum_{i=1}^{n}(i+1) && \text{Factor constant } 1/n^2 \text{ out of sum.} \\ &= \frac{1}{n^2}\left(\sum_{i=1}^{n} i + \sum_{i=1}^{n} 1\right) && \text{Write as two sums.} \\ &= \frac{1}{n^2}\left[\frac{n(n+1)}{2} + n\right] && \text{Apply Theorem 4.2.} \\ &= \frac{1}{n^2}\left[\frac{n^2+3n}{2}\right] && \text{Simplify.} \\ &= \frac{n+3}{2n}. && \text{Simplify.} \end{aligned}$$

Now you can evaluate the sum by substituting the appropriate values of n, as shown in the table at the left.

n	$\displaystyle\sum_{i=1}^{n} \frac{i+1}{n^2} = \frac{n+3}{2n}$
10	0.65000
100	0.51500
1,000	0.50150
10,000	0.50015

Try It **Exploration A** **Exploration B**

In the table, note that the sum appears to approach a limit as n increases. Although the discussion of limits at infinity in Section 3.5 applies to a variable x, where x can be any real number, many of the same results hold true for limits involving the variable n, where n is restricted to positive integer values. So, to find the limit of $(n+3)/2n$ as n approaches infinity, you can write

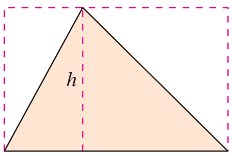
$$\lim_{n\to\infty} \frac{n+3}{2n} = \frac{1}{2}.$$

Area

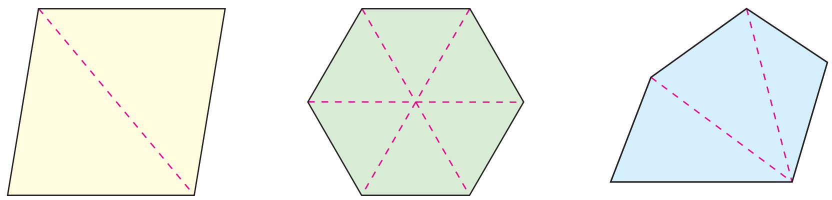
Rectangle: $A = bh$
Figure 4.5

In Euclidean geometry, the simplest type of plane region is a rectangle. Although people often say that the *formula* for the area of a rectangle is $A = bh$, as shown in Figure 4.5, it is actually more proper to say that this is the *definition* of the **area of a rectangle.**

From this definition, you can develop formulas for the areas of many other plane regions. For example, to determine the area of a triangle, you can form a rectangle whose area is twice that of the triangle, as shown in Figure 4.6. Once you know how to find the area of a triangle, you can determine the area of any polygon by subdividing the polygon into triangular regions, as shown in Figure 4.7.



Triangle: $A = \frac{1}{2}bh$
Figure 4.6



Parallelogram Hexagon Polygon
Figure 4.7

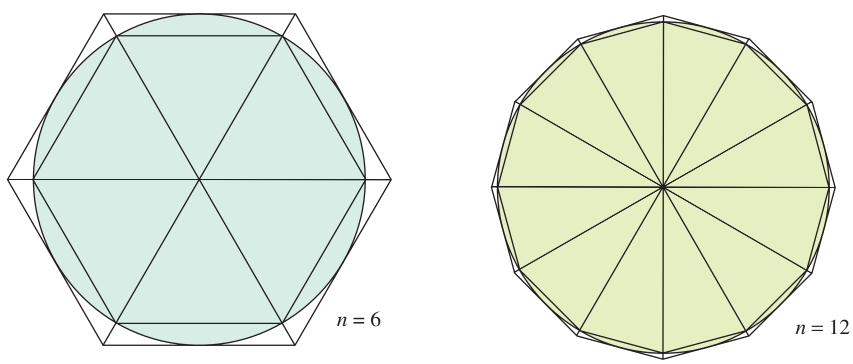
ARCHIMEDES (287–212 B.C.)

Archimedes used the method of exhaustion to derive formulas for the areas of ellipses, parabolic segments, and sectors of a spiral. He is considered to have been the greatest applied mathematician of antiquity.

MathBio

Finding the areas of regions other than polygons is more difficult. The ancient Greeks were able to determine formulas for the areas of some general regions (principally those bounded by conics) by the *exhaustion* method. The clearest description of this method was given by Archimedes. Essentially, the method is a limiting process in which the area is squeezed between two polygons—one inscribed in the region and one circumscribed about the region.

For instance, in Figure 4.8 the area of a circular region is approximated by an n-sided inscribed polygon and an n-sided circumscribed polygon. For each value of n the area of the inscribed polygon is less than the area of the circle, and the area of the circumscribed polygon is greater than the area of the circle. Moreover, as n increases, the areas of both polygons become better and better approximations of the area of the circle.



The exhaustion method for finding the area of a circular region
Figure 4.8

Animation

FOR FURTHER INFORMATION For an alternative development of the formula for the area of a circle, see the article "Proof Without Words: Area of a Disk is πR^2" by Russell Jay Hendel in *Mathematics Magazine*.

MathArticle

A process that is similar to that used by Archimedes to determine the area of a plane region is used in the remaining examples in this section.

The Area of a Plane Region

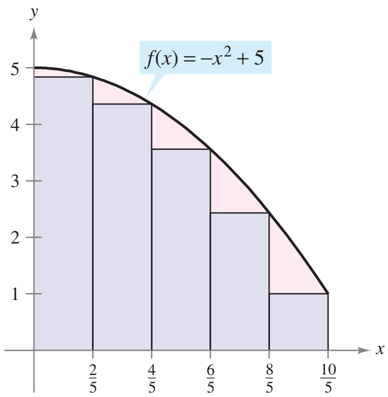
Recall from Section 1.1 that the origins of calculus are connected to two classic problems: the tangent line problem and the area problem. Example 3 begins the investigation of the area problem.

EXAMPLE 3 Approximating the Area of a Plane Region

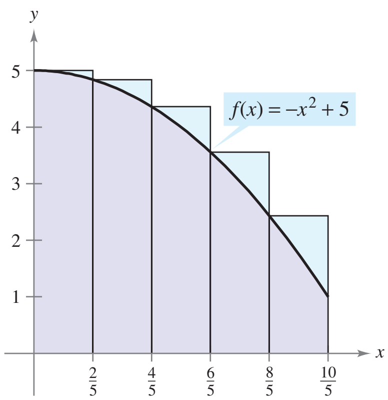
Use the five rectangles in Figure 4.9(a) and (b) to find *two* approximations of the area of the region lying between the graph of

$$f(x) = -x^2 + 5$$

and the x-axis between $x = 0$ and $x = 2$.



(a) The area of the parabolic region is greater than the area of the rectangles.



(b) The area of the parabolic region is less than the area of the rectangles.

Figure 4.9

Solution

a. The right endpoints of the five intervals are $\frac{2}{5}i$, where $i = 1, 2, 3, 4, 5$. The width of each rectangle is $\frac{2}{5}$, and the height of each rectangle can be obtained by evaluating f at the right endpoint of each interval.

$$\left[0, \frac{2}{5}\right], \left[\frac{2}{5}, \frac{4}{5}\right], \left[\frac{4}{5}, \frac{6}{5}\right], \left[\frac{6}{5}, \frac{8}{5}\right], \left[\frac{8}{5}, \frac{10}{5}\right]$$

Evaluate f at the right endpoints of these intervals.

The sum of the areas of the five rectangles is

$$\sum_{i=1}^{5} \underbrace{f\left(\frac{2i}{5}\right)}_{\text{Height}} \underbrace{\left(\frac{2}{5}\right)}_{\text{Width}} = \sum_{i=1}^{5}\left[-\left(\frac{2i}{5}\right)^2 + 5\right]\left(\frac{2}{5}\right) = \frac{162}{25} = 6.48.$$

Because each of the five rectangles lies inside the parabolic region, you can conclude that the area of the parabolic region is greater than 6.48.

b. The left endpoints of the five intervals are $\frac{2}{5}(i - 1)$, where $i = 1, 2, 3, 4, 5$. The width of each rectangle is $\frac{2}{5}$, and the height of each rectangle can be obtained by evaluating f at the left endpoint of each interval.

$$\sum_{i=1}^{5} \underbrace{f\left(\frac{2i-2}{5}\right)}_{\text{Height}} \underbrace{\left(\frac{2}{5}\right)}_{\text{Width}} = \sum_{i=1}^{5}\left[-\left(\frac{2i-2}{5}\right)^2 + 5\right]\left(\frac{2}{5}\right) = \frac{202}{25} = 8.08.$$

Because the parabolic region lies within the union of the five rectangular regions, you can conclude that the area of the parabolic region is less than 8.08.

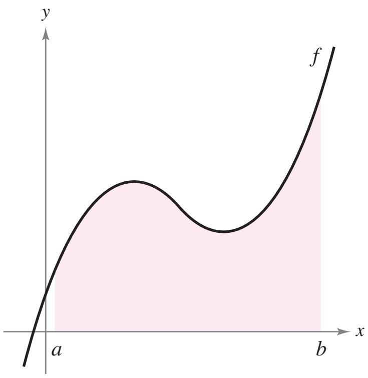
By combining the results in parts (a) and (b), you can conclude that

$$6.48 < (\text{Area of region}) < 8.08.$$

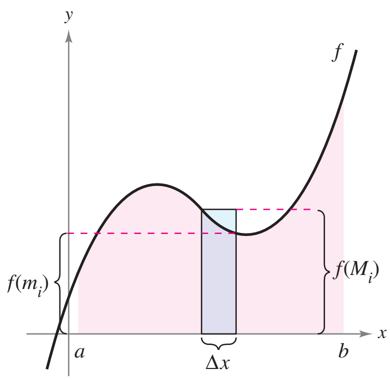
Try It Exploration A Exploration B Video

NOTE By increasing the number of rectangles used in Example 3, you can obtain closer and closer approximations of the area of the region. For instance, using 25 rectangles of width $\frac{2}{25}$ each, you can conclude that

$$7.17 < (\text{Area of region}) < 7.49.$$



The region under a curve
Figure 4.10



The interval $[a, b]$ is divided into n subintervals of width $\Delta x = \dfrac{b-a}{n}$.
Figure 4.11

Upper and Lower Sums

The procedure used in Example 3 can be generalized as follows. Consider a plane region bounded above by the graph of a nonnegative, continuous function $y = f(x)$, as shown in Figure 4.10. The region is bounded below by the x-axis, and the left and right boundaries of the region are the vertical lines $x = a$ and $x = b$.

To approximate the area of the region, begin by subdividing the interval $[a, b]$ into n subintervals, each of width $\Delta x = (b - a)/n$, as shown in Figure 4.11. The endpoints of the intervals are as follows.

$$\overbrace{a + 0(\Delta x)}^{a = x_0} < \overbrace{a + 1(\Delta x)}^{x_1} < \overbrace{a + 2(\Delta x)}^{x_2} < \cdots < \overbrace{a + n(\Delta x)}^{x_n = b}$$

Because f is continuous, the Extreme Value Theorem guarantees the existence of a minimum and a maximum value of $f(x)$ in *each* subinterval.

$f(m_i)$ = Minimum value of $f(x)$ in ith subinterval

$f(M_i)$ = Maximum value of $f(x)$ in ith subinterval

Next, define an **inscribed rectangle** lying *inside* the ith subregion and a **circumscribed rectangle** extending *outside* the ith subregion. The height of the ith inscribed rectangle is $f(m_i)$ and the height of the ith circumscribed rectangle is $f(M_i)$. For *each* i, the area of the inscribed rectangle is less than or equal to the area of the circumscribed rectangle.

$$\begin{pmatrix}\text{Area of inscribed}\\ \text{rectangle}\end{pmatrix} = f(m_i)\,\Delta x \le f(M_i)\,\Delta x = \begin{pmatrix}\text{Area of circumscribed}\\ \text{rectangle}\end{pmatrix}$$

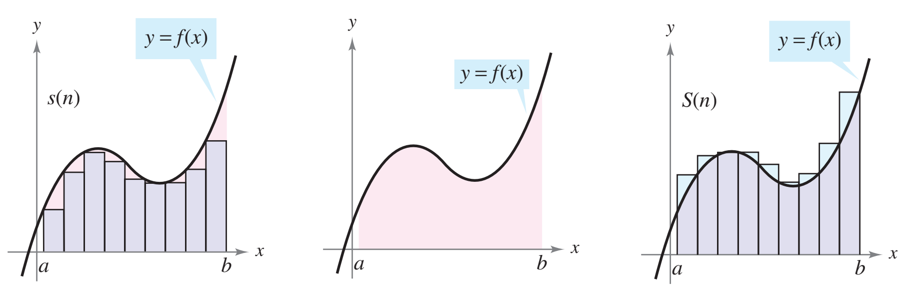
The sum of the areas of the inscribed rectangles is called a **lower sum,** and the sum of the areas of the circumscribed rectangles is called an **upper sum.**

$$\text{Lower sum} = s(n) = \sum_{i=1}^{n} f(m_i)\,\Delta x \qquad \text{Area of inscribed rectangles}$$

$$\text{Upper sum} = S(n) = \sum_{i=1}^{n} f(M_i)\,\Delta x \qquad \text{Area of circumscribed rectangles}$$

From Figure 4.12, you can see that the lower sum $s(n)$ is less than or equal to the upper sum $S(n)$. Moreover, the actual area of the region lies between these two sums.

$$s(n) \le (\text{Area of region}) \le S(n)$$



Area of inscribed rectangles is less than area of region.

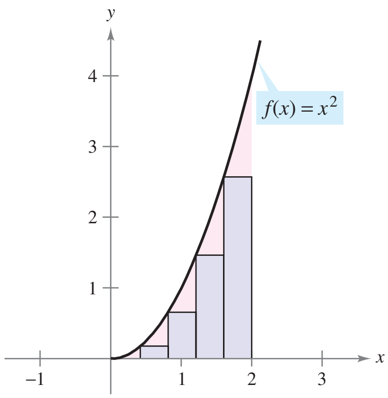
Area of region

Area of circumscribed rectangles is greater than area of region.

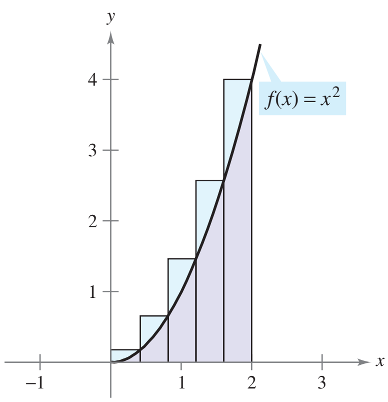
Figure 4.12

As n increases, both the lower sum $s(n)$ and the upper sum $S(n)$ become closer to the actual area of the region. View the animation to see this.

Animation



Inscribed rectangles



Circumscribed rectangles
Figure 4.13

Editable Graph

EXAMPLE 4 Finding Upper and Lower Sums for a Region

Find the upper and lower sums for the region bounded by the graph of $f(x) = x^2$ and the x-axis between $x = 0$ and $x = 2$.

Solution To begin, partition the interval $[0, 2]$ into n subintervals, each of width

$$\Delta x = \frac{b-a}{n} = \frac{2-0}{n} = \frac{2}{n}.$$

Figure 4.13 shows the endpoints of the subintervals and several inscribed and circumscribed rectangles. Because f is increasing on the interval $[0, 2]$, the minimum value on each subinterval occurs at the left endpoint, and the maximum value occurs at the right endpoint.

Left Endpoints

$$m_i = 0 + (i-1)\left(\frac{2}{n}\right) = \frac{2(i-1)}{n}$$

Right Endpoints

$$M_i = 0 + i\left(\frac{2}{n}\right) = \frac{2i}{n}$$

Using the left endpoints, the lower sum is

$$\begin{aligned}
s(n) = \sum_{i=1}^{n} f(m_i)\,\Delta x &= \sum_{i=1}^{n} f\left[\frac{2(i-1)}{n}\right]\left(\frac{2}{n}\right)\\
&= \sum_{i=1}^{n}\left[\frac{2(i-1)}{n}\right]^2\left(\frac{2}{n}\right)\\
&= \sum_{i=1}^{n}\left(\frac{8}{n^3}\right)(i^2 - 2i + 1)\\
&= \frac{8}{n^3}\left(\sum_{i=1}^{n} i^2 - 2\sum_{i=1}^{n} i + \sum_{i=1}^{n} 1\right)\\
&= \frac{8}{n^3}\left\{\frac{n(n+1)(2n+1)}{6} - 2\left[\frac{n(n+1)}{2}\right] + n\right\}\\
&= \frac{4}{3n^3}(2n^3 - 3n^2 + n)\\
&= \frac{8}{3} - \frac{4}{n} + \frac{4}{3n^2}. \qquad \text{Lower sum}
\end{aligned}$$

Using the right endpoints, the upper sum is

$$\begin{aligned}
S(n) = \sum_{i=1}^{n} f(M_i)\,\Delta x &= \sum_{i=1}^{n} f\left(\frac{2i}{n}\right)\left(\frac{2}{n}\right)\\
&= \sum_{i=1}^{n}\left(\frac{2i}{n}\right)^2\left(\frac{2}{n}\right)\\
&= \sum_{i=1}^{n}\left(\frac{8}{n^3}\right)i^2\\
&= \frac{8}{n^3}\left[\frac{n(n+1)(2n+1)}{6}\right]\\
&= \frac{4}{3n^3}(2n^3 + 3n^2 + n)\\
&= \frac{8}{3} + \frac{4}{n} + \frac{4}{3n^2}. \qquad \text{Upper sum}
\end{aligned}$$

Try It   Exploration A   Exploration B

EXPLORATION

For the region given in Example 4, evaluate the lower sum

$$s(n) = \frac{8}{3} - \frac{4}{n} + \frac{4}{3n^2}$$

and the upper sum

$$S(n) = \frac{8}{3} + \frac{4}{n} + \frac{4}{3n^2}$$

for $n = 10, 100$, and 1000. Use your results to determine the area of the region.

Example 4 illustrates some important things about lower and upper sums. First, notice that for any value of n, the lower sum is less than (or equal to) the upper sum.

$$s(n) = \frac{8}{3} - \frac{4}{n} + \frac{4}{3n^2} < \frac{8}{3} + \frac{4}{n} + \frac{4}{3n^2} = S(n)$$

Second, the difference between these two sums lessens as n increases. In fact, if you take the limits as $n \to \infty$, both the upper sum and the lower sum approach $\frac{8}{3}$.

$$\lim_{n\to\infty} s(n) = \lim_{n\to\infty}\left(\frac{8}{3} - \frac{4}{n} + \frac{4}{3n^2}\right) = \frac{8}{3} \qquad \text{Lower sum limit}$$

$$\lim_{n\to\infty} S(n) = \lim_{n\to\infty}\left(\frac{8}{3} + \frac{4}{n} + \frac{4}{3n^2}\right) = \frac{8}{3} \qquad \text{Upper sum limit}$$

The next theorem shows that the equivalence of the limits (as $n \to \infty$) of the upper and lower sums is not mere coincidence. It is true for all functions that are continuous and nonnegative on the closed interval $[a, b]$. The proof of this theorem is best left to a course in advanced calculus.

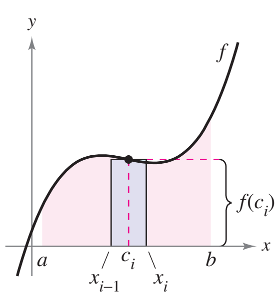
THEOREM 4.3 Limits of the Lower and Upper Sums

Let f be continuous and nonnegative on the interval $[a, b]$. The limits as $n \to \infty$ of both the lower and upper sums exist and are equal to each other. That is,

$$\begin{aligned}\lim_{n\to\infty} s(n) &= \lim_{n\to\infty} \sum_{i=1}^{n} f(m_i)\,\Delta x \\ &= \lim_{n\to\infty} \sum_{i=1}^{n} f(M_i)\,\Delta x \\ &= \lim_{n\to\infty} S(n)\end{aligned}$$

where $\Delta x = (b - a)/n$ and $f(m_i)$ and $f(M_i)$ are the minimum and maximum values of f on the subinterval.

Because the same limit is attained for both the minimum value $f(m_i)$ and the maximum value $f(M_i)$, it follows from the Squeeze Theorem (Theorem 1.8) that the choice of x in the ith subinterval does not affect the limit. This means that you are free to choose an *arbitrary* x-value in the ith subinterval, as in the following *definition of the area of a region in the plane*.



The width of the ith subinterval is $\Delta x = x_i - x_{i-1}$.
Figure 4.14

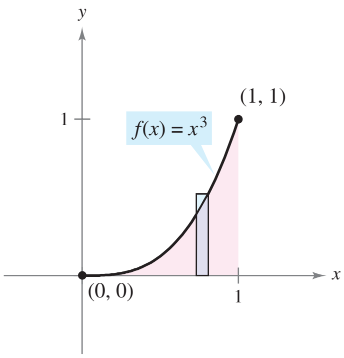
Definition of the Area of a Region in the Plane

Let f be continuous and nonnegative on the interval $[a, b]$. The area of the region bounded by the graph of f, the x-axis, and the vertical lines $x = a$ and $x = b$ is

$$\text{Area} = \lim_{n\to\infty} \sum_{i=1}^{n} f(c_i)\,\Delta x, \qquad x_{i-1} \le c_i \le x_i$$

where $\Delta x = (b - a)/n$ (see Figure 4.14).

Video



The area of the region bounded by the graph of f, the x-axis, $x = 0$, and $x = 1$ is $\frac{1}{4}$.
Figure 4.15

Editable Graph

EXAMPLE 5 Finding Area by the Limit Definition

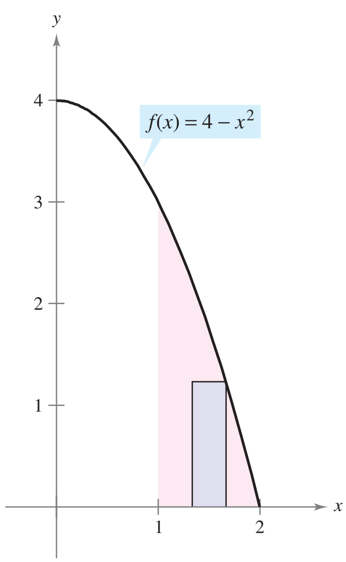
Find the area of the region bounded by the graph $f(x) = x^3$, the x-axis, and the vertical lines $x = 0$ and $x = 1$, as shown in Figure 4.15.

Solution Begin by noting that f is continuous and nonnegative on the interval $[0, 1]$. Next, partition the interval $[0, 1]$ into n subintervals, each of width $\Delta x = 1/n$. According to the definition of area, you can choose any x-value in the ith subinterval. For this example, the right endpoints $c_i = i/n$ are convenient.

$$\begin{aligned}
\text{Area} = \lim_{n\to\infty} \sum_{i=1}^{n} f(c_i)\,\Delta x &= \lim_{n\to\infty} \sum_{i=1}^{n} \left(\frac{i}{n}\right)^3 \left(\frac{1}{n}\right) && \text{Right endpoints: } c_i = \frac{i}{n} \\
&= \lim_{n\to\infty} \frac{1}{n^4} \sum_{i=1}^{n} i^3 \\
&= \lim_{n\to\infty} \frac{1}{n^4} \left[\frac{n^2(n+1)^2}{4}\right] \\
&= \lim_{n\to\infty} \left(\frac{1}{4} + \frac{1}{2n} + \frac{1}{4n^2}\right) \\
&= \frac{1}{4}
\end{aligned}$$

The area of the region is $\frac{1}{4}$.

Try It Exploration A



The area of the region bounded by the graph of f, the x-axis, $x = 1$, and $x = 2$ is $\frac{5}{3}$.
Figure 4.16

Editable Graph

EXAMPLE 6 Finding Area by the Limit Definition

Find the area of the region bounded by the graph of $f(x) = 4 - x^2$, the x-axis, and the vertical lines $x = 1$ and $x = 2$, as shown in Figure 4.16.

Solution The function f is continuous and nonnegative on the interval $[1, 2]$, and so begin by partitioning the interval into n subintervals, each of width $\Delta x = 1/n$. Choosing the right endpoint

$$c_i = a + i\Delta x = 1 + \frac{i}{n} \qquad \text{Right endpoints}$$

of each subinterval, you obtain

$$\begin{aligned}
\text{Area} = \lim_{n\to\infty} \sum_{i=1}^{n} f(c_i)\,\Delta x &= \lim_{n\to\infty} \sum_{i=1}^{n} \left[4 - \left(1 + \frac{i}{n}\right)^2\right]\left(\frac{1}{n}\right) \\
&= \lim_{n\to\infty} \sum_{i=1}^{n} \left(3 - \frac{2i}{n} - \frac{i^2}{n^2}\right)\left(\frac{1}{n}\right) \\
&= \lim_{n\to\infty} \left(\frac{1}{n}\sum_{i=1}^{n} 3 - \frac{2}{n^2}\sum_{i=1}^{n} i - \frac{1}{n^3}\sum_{i=1}^{n} i^2\right) \\
&= \lim_{n\to\infty} \left[3 - \left(1 + \frac{1}{n}\right) - \left(\frac{1}{3} + \frac{1}{2n} + \frac{1}{6n^2}\right)\right] \\
&= 3 - 1 - \frac{1}{3} \\
&= \frac{5}{3}.
\end{aligned}$$

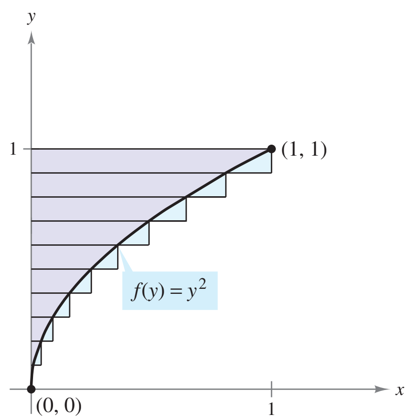
The area of the region is $\frac{5}{3}$.

Try It Exploration A Open Exploration

The last example in this section looks at a region that is bounded by the y-axis (rather than by the x-axis).

EXAMPLE 7 A Region Bounded by the y-axis

Find the area of the region bounded by the graph of $f(y) = y^2$ and the y-axis for $0 \le y \le 1$, as shown in Figure 4.17.



The area of the region bounded by the graph of f and the y-axis for $0 \le y \le 1$ is $\frac{1}{3}$.
Figure 4.17

Editable Graph

Solution When f is a continuous, nonnegative function of y, you still can use the same basic procedure shown in Examples 5 and 6. Begin by partitioning the interval $[0, 1]$ into n subintervals, each of width $\Delta y = 1/n$. Then, using the upper endpoints $c_i = i/n$, you obtain

$$\begin{aligned}
\text{Area} &= \lim_{n\to\infty} \sum_{i=1}^{n} f(c_i)\,\Delta y = \lim_{n\to\infty} \sum_{i=1}^{n} \left(\frac{i}{n}\right)^2 \left(\frac{1}{n}\right) && \text{Upper endpoints: } c_i = \frac{i}{n} \\
&= \lim_{n\to\infty} \frac{1}{n^3} \sum_{i=1}^{n} i^2 \\
&= \lim_{n\to\infty} \frac{1}{n^3} \left[\frac{n(n+1)(2n+1)}{6}\right] \\
&= \lim_{n\to\infty} \left(\frac{1}{3} + \frac{1}{2n} + \frac{1}{6n^2}\right) \\
&= \frac{1}{3}.
\end{aligned}$$

The area of the region is $\frac{1}{3}$.

Try It

Exploration A

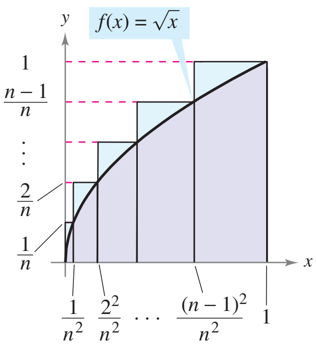
Section 4.3

Riemann Sums and Definite Integrals

- **Understand the definition of a Riemann sum.**
- **Evaluate a definite integral using limits.**
- **Evaluate a definite integral using properties of definite integrals.**

Riemann Sums

In the definition of area given in Section 4.2, the partitions have subintervals of *equal width*. This was done only for computational convenience. The following example shows that it is not necessary to have subintervals of equal width.



The subintervals do not have equal widths.
Figure 4.18

Editable Graph

EXAMPLE 1 A Partition with Subintervals of Unequal Widths

Consider the region bounded by the graph of $f(x) = \sqrt{x}$ and the x-axis for $0 \le x \le 1$, as shown in Figure 4.18. Evaluate the limit

$$\lim_{n\to\infty} \sum_{i=1}^{n} f(c_i)\, \Delta x_i$$

where c_i is the right endpoint of the partition given by $c_i = i^2/n^2$ and Δx_i is the width of the ith interval.

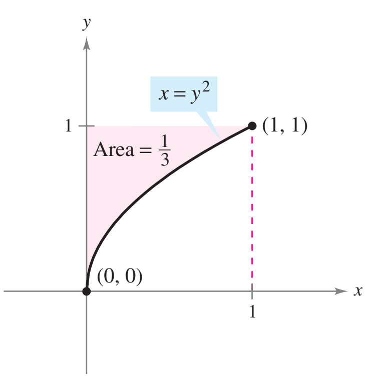
Solution The width of the ith interval is given by

$$\begin{aligned} \Delta x_i &= \frac{i^2}{n^2} - \frac{(i-1)^2}{n^2} \\ &= \frac{i^2 - i^2 + 2i - 1}{n^2} \\ &= \frac{2i-1}{n^2}. \end{aligned}$$

So, the limit is

$$\begin{aligned} \lim_{n\to\infty} \sum_{i=1}^{n} f(c_i)\, \Delta x_i &= \lim_{n\to\infty} \sum_{i=1}^{n} \sqrt{\frac{i^2}{n^2}}\left(\frac{2i-1}{n^2}\right) \\ &= \lim_{n\to\infty} \frac{1}{n^3} \sum_{i=1}^{n} (2i^2 - i) \\ &= \lim_{n\to\infty} \frac{1}{n^3}\left[2\left(\frac{n(n+1)(2n+1)}{6}\right) - \frac{n(n+1)}{2}\right] \\ &= \lim_{n\to\infty} \frac{4n^3 + 3n^2 - n}{6n^3} \\ &= \frac{2}{3}. \end{aligned}$$

Try It   Exploration A



The area of the region bounded by the graph of $x = y^2$ and the y-axis for $0 \le y \le 1$ is $\frac{1}{3}$.
Figure 4.19

From Example 7 in Section 4.2, you know that the region shown in Figure 4.19 has an area of $\frac{1}{3}$. Because the square bounded by $0 \le x \le 1$ and $0 \le y \le 1$ has an area of 1, you can conclude that the area of the region shown in Figure 4.18 has an area of $\frac{2}{3}$. This agrees with the limit found in Example 1, even though that example used a partition having subintervals of unequal widths. The reason this particular partition gave the proper area is that as n increases, the *width of the largest subinterval approaches zero*. This is a key feature of the development of definite integrals.

Georg Friedrich Bernhard Riemann (1826–1866)

German mathematician Riemann did his most famous work in the areas of non-Euclidean geometry, differential equations, and number theory. It was Riemann's results in physics and mathematics that formed the structure on which Einstein's theory of general relativity is based.

MathBio

In the preceding section, the limit of a sum was used to define the area of a region in the plane. Finding area by this means is only one of *many* applications involving the limit of a sum. A similar approach can be used to determine quantities as diverse as arc lengths, average values, centroids, volumes, work, and surface areas. The following definition is named after Georg Friedrich Bernhard Riemann. Although the definite integral had been defined and used long before the time of Riemann, he generalized the concept to cover a broader category of functions.

In the following definition of a Riemann sum, note that the function f has no restrictions other than being defined on the interval $[a, b]$. (In the preceding section, the function f was assumed to be continuous and nonnegative because we were dealing with the area under a curve.)

Definition of a Riemann Sum

Let f be defined on the closed interval $[a, b]$, and let Δ be a partition of $[a, b]$ given by

$$a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = b$$

where Δx_i is the width of the ith subinterval. If c_i is *any* point in the ith subinterval, then the sum

$$\sum_{i=1}^{n} f(c_i)\,\Delta x_i, \qquad x_{i-1} \le c_i \le x_i$$

is called a **Riemann sum** of f for the partition Δ.

Video

NOTE The sums in Section 4.2 are examples of Riemann sums, but there are more general Riemann sums than those covered there.

The width of the largest subinterval of a partition Δ is the **norm** of the partition and is denoted by $\|\Delta\|$. If every subinterval is of equal width, the partition is **regular** and the norm is denoted by

$$\|\Delta\| = \Delta x = \frac{b-a}{n}. \qquad \text{Regular partition}$$

For a general partition, the norm is related to the number of subintervals of $[a, b]$ in the following way.

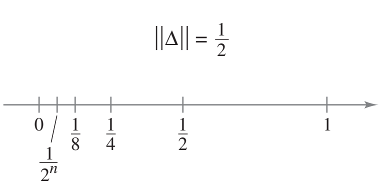
$$\frac{b-a}{\|\Delta\|} \le n \qquad \text{General partition}$$

So, the number of subintervals in a partition approaches infinity as the norm of the partition approaches 0. That is, $\|\Delta\| \to 0$ implies that $n \to \infty$.

The converse of this statement is not true. For example, let Δ_n be the partition of the interval $[0, 1]$ given by

$$0 < \frac{1}{2^n} < \frac{1}{2^{n-1}} < \cdots < \frac{1}{8} < \frac{1}{4} < \frac{1}{2} < 1.$$

As shown in Figure 4.20, for any positive value of n, the norm of the partition Δ_n is $\frac{1}{2}$. So, letting n approach infinity does not force $\|\Delta\|$ to approach 0. In a regular partition, however, the statements $\|\Delta\| \to 0$ and $n \to \infty$ are equivalent.



$n \to \infty$ does not imply that $\|\Delta\| \to 0$.
Figure 4.20

Definite Integrals

To define the definite integral, consider the following limit.

$$\lim_{\|\Delta\| \to 0} \sum_{i=1}^{n} f(c_i)\,\Delta x_i = L$$

To say that this limit exists means that for $\varepsilon > 0$ there exists a $\delta > 0$ such that for every partition with $\|\Delta\| < \delta$ it follows that

$$\left| L - \sum_{i=1}^{n} f(c_i)\,\Delta x_i \right| < \varepsilon.$$

(This must be true for any choice of c_i in the ith subinterval of Δ.)

FOR FURTHER INFORMATION For insight into the history of the definite integral, see the article "The Evolution of Integration" by A. Shenitzer and J. Steprāns in The *American Mathematical Monthly*.

MathArticle

Definition of a Definite Integral

If f is defined on the closed interval $[a, b]$ and the limit

$$\lim_{\|\Delta\| \to 0} \sum_{i=1}^{n} f(c_i)\,\Delta x_i$$

exists (as described above), then f is **integrable** on $[a, b]$ and the limit is denoted by

$$\lim_{\|\Delta\| \to 0} \sum_{i=1}^{n} f(c_i)\,\Delta x_i = \int_a^b f(x)\,dx.$$

The limit is called the **definite integral** of f from a to b. The number a is the **lower limit** of integration, and the number b is the **upper limit** of integration.

It is not a coincidence that the notation for definite integrals is similar to that used for indefinite integrals. You will see why in the next section when the Fundamental Theorem of Calculus is introduced. For now it is important to see that definite integrals and indefinite integrals are different identities. A definite integral is a *number*, whereas an indefinite integral is a *family of functions*.

A sufficient condition for a function f to be integrable on $[a, b]$ is that it is continuous on $[a, b]$. A proof of this theorem is beyond the scope of this text.

THEOREM 4.4 Continuity Implies Integrability

If a function f is continuous on the closed interval $[a, b]$, then f is integrable on $[a, b]$.

EXPLORATION

The Converse of Theorem 4.4 Is the converse of Theorem 4.4 true? That is, if a function is integrable, does it have to be continuous? Explain your reasoning and give examples.

Describe the relationships among continuity, differentiability, and integrability. Which is the strongest condition? Which is the weakest? Which conditions imply other conditions?

EXAMPLE 2 Evaluating a Definite Integral as a Limit

Evaluate the definite integral $\int_{-2}^{1} 2x\,dx$.

Solution The function $f(x) = 2x$ is integrable on the interval $[-2, 1]$ because it is continuous on $[-2, 1]$. Moreover, the definition of integrability implies that any partition whose norm approaches 0 can be used to determine the limit. For computational convenience, define Δ by subdividing $[-2, 1]$ into n subintervals of equal width

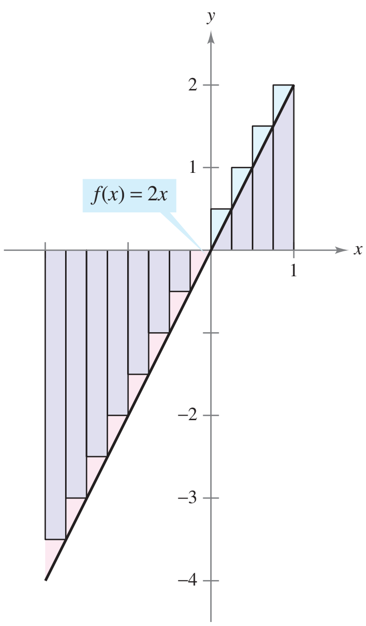
$$\Delta x_i = \Delta x = \frac{b-a}{n} = \frac{3}{n}.$$

Choosing c_i as the right endpoint of each subinterval produces

$$c_i = a + i(\Delta x) = -2 + \frac{3i}{n}.$$

So, the definite integral is given by

$$\begin{aligned}
\int_{-2}^{1} 2x\,dx &= \lim_{\|\Delta\|\to 0} \sum_{i=1}^{n} f(c_i)\,\Delta x_i \\
&= \lim_{n\to\infty} \sum_{i=1}^{n} f(c_i)\,\Delta x \\
&= \lim_{n\to\infty} \sum_{i=1}^{n} 2\left(-2 + \frac{3i}{n}\right)\left(\frac{3}{n}\right) \\
&= \lim_{n\to\infty} \frac{6}{n} \sum_{i=1}^{n} \left(-2 + \frac{3i}{n}\right) \\
&= \lim_{n\to\infty} \frac{6}{n} \left\{-2n + \frac{3}{n}\left[\frac{n(n+1)}{2}\right]\right\} \\
&= \lim_{n\to\infty} \left(-12 + 9 + \frac{9}{n}\right) \\
&= -3.
\end{aligned}$$

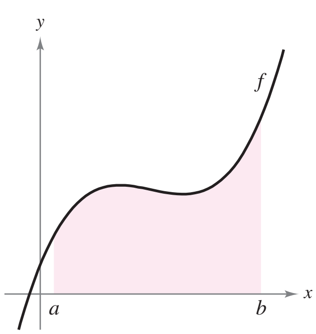


Because the definite integral is negative, it does not represent the area of the region.
Figure 4.21

Editable Graph

Try It | Exploration A | Video | Video

Because the definite integral in Example 2 is negative, it *does not* represent the area of the region shown in Figure 4.21. Definite integrals can be positive, negative, or zero. For a definite integral to be interpreted as an area (as defined in Section 4.2), the function f must be continuous and nonnegative on $[a, b]$, as stated in the following theorem. (The proof of this theorem is straightforward—you simply use the definition of area given in Section 4.2.)



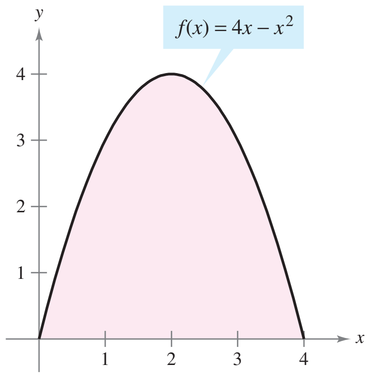
You can use a definite integral to find the area of the region bounded by the graph of f, the x-axis, $x = a$, and $x = b$.
Figure 4.22

THEOREM 4.5 The Definite Integral as the Area of a Region

If f is continuous and nonnegative on the closed interval $[a, b]$, then the area of the region bounded by the graph of f, the x-axis, and the vertical lines $x = a$ and $x = b$ is given by

$$\text{Area} = \int_a^b f(x)\,dx.$$

(See Figure 4.22.)



$\text{Area} = \int_0^4 (4x - x^2)\,dx$

Figure 4.23

As an example of Theorem 4.5, consider the region bounded by the graph of

$$f(x) = 4x - x^2$$

and the x-axis, as shown in Figure 4.23. Because f is continuous and nonnegative on the closed interval $[0, 4]$, the area of the region is

$$\text{Area} = \int_0^4 (4x - x^2)\,dx.$$

A straightforward technique for evaluating a definite integral such as this will be discussed in Section 4.4. For now, however, you can evaluate a definite integral in two ways—you can use the limit definition *or* you can check to see whether the definite integral represents the area of a common geometric region such as a rectangle, triangle, or semicircle.

EXAMPLE 3 Areas of Common Geometric Figures

Sketch the region corresponding to each definite integral. Then evaluate each integral using a geometric formula.

a. $\int_1^3 4\,dx$ **b.** $\int_0^3 (x + 2)\,dx$ **c.** $\int_{-2}^2 \sqrt{4 - x^2}\,dx$

Solution A sketch of each region is shown in Figure 4.24.

a. This region is a rectangle of height 4 and width 2.

$$\int_1^3 4\,dx = (\text{Area of rectangle}) = 4(2) = 8$$

b. This region is a trapezoid with an altitude of 3 and parallel bases of lengths 2 and 5. The formula for the area of a trapezoid is $\frac{1}{2}h(b_1 + b_2)$.

$$\int_0^3 (x + 2)\,dx = (\text{Area of trapezoid}) = \frac{1}{2}(3)(2 + 5) = \frac{21}{2}$$

c. This region is a semicircle of radius 2. The formula for the area of a semicircle is $\frac{1}{2}\pi r^2$.

$$\int_{-2}^2 \sqrt{4 - x^2}\,dx = (\text{Area of semicircle}) = \frac{1}{2}\pi(2^2) = 2\pi$$

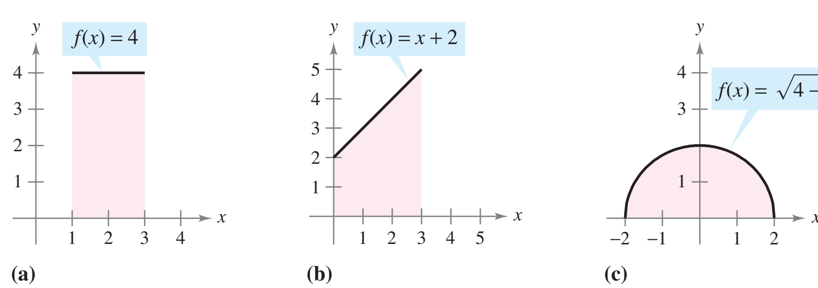
NOTE The variable of integration in a definite integral is sometimes called a *dummy variable* because it can be replaced by any other variable without changing the value of the integral. For instance, the definite integrals

$$\int_0^3 (x + 2)\,dx$$

and

$$\int_0^3 (t + 2)\,dt$$

have the same value.



Editable Graph Editable Graph Editable Graph

Figure 4.24

Try It Exploration A

Properties of Definite Integrals

The definition of the definite integral of f on the interval $[a, b]$ specifies that $a < b$. Now, however, it is convenient to extend the definition to cover cases in which $a = b$ or $a > b$. Geometrically, the following two definitions seem reasonable. For instance, it makes sense to define the area of a region of zero width and finite height to be 0.

Definitions of Two Special Definite Integrals

1. If f is defined at $x = a$, then we define $\int_a^a f(x)\,dx = 0$.

2. If f is integrable on $[a, b]$, then we define $\int_b^a f(x)\,dx = -\int_a^b f(x)\,dx$.

EXAMPLE 4 Evaluating Definite Integrals

a. Because the sine function is defined at $x = \pi$, and the upper and lower limits of integration are equal, you can write

$$\int_\pi^\pi \sin x\,dx = 0.$$

b. The integral $\int_3^0 (x + 2)\,dx$ is the same as that given in Example 3(b) except that the upper and lower limits are interchanged. Because the integral in Example 3(b) has a value of $\frac{21}{2}$, you can write

$$\int_3^0 (x + 2)\,dx = -\int_0^3 (x + 2)\,dx = -\frac{21}{2}.$$

The editable graph feature below allows you to edit the graph of a function.

Editable Graph

Try It | Exploration A | Open Exploration

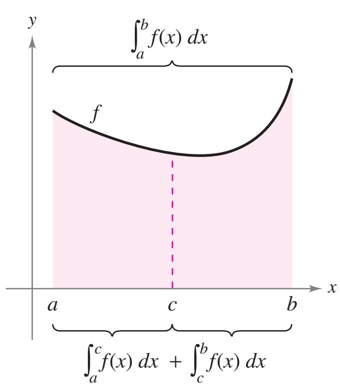


Figure 4.25

In Figure 4.25, the larger region can be divided at $x = c$ into two subregions whose intersection is a line segment. Because the line segment has zero area, it follows that the area of the larger region is equal to the sum of the areas of the two smaller regions.

THEOREM 4.6 Additive Interval Property

If f is integrable on the three closed intervals determined by a, b, and c, then

$$\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx.$$

EXAMPLE 5 Using the Additive Interval Property

$$\int_{-1}^1 |x|\,dx = \int_{-1}^0 -x\,dx + \int_0^1 x\,dx \quad \text{Theorem 4.6}$$

$$= \frac{1}{2} + \frac{1}{2} \quad \text{Area of a triangle}$$

$$= 1$$

Try It | Exploration A

Because the definite integral is defined as the limit of a sum, it inherits the properties of summation given at the top of page 260.

THEOREM 4.7 Properties of Definite Integrals

If f and g are integrable on $[a, b]$ and k is a constant, then the functions of kf and $f \pm g$ are integrable on $[a, b]$, and

1. $\displaystyle\int_a^b kf(x)\,dx = k\int_a^b f(x)\,dx$

2. $\displaystyle\int_a^b [f(x) \pm g(x)]\,dx = \int_a^b f(x)\,dx \pm \int_a^b g(x)\,dx.$

Note that Property 2 of Theorem 4.7 can be extended to cover any finite number of functions. For example,

$$\int_a^b [f(x) + g(x) + h(x)]\,dx = \int_a^b f(x)\,dx + \int_a^b g(x)\,dx + \int_a^b h(x)\,dx.$$

EXAMPLE 6 Evaluation of a Definite Integral

Evaluate $\displaystyle\int_1^3 (-x^2 + 4x - 3)\,dx$ using each of the following values.

$$\int_1^3 x^2\,dx = \frac{26}{3}, \qquad \int_1^3 x\,dx = 4, \qquad \int_1^3 dx = 2$$

Solution

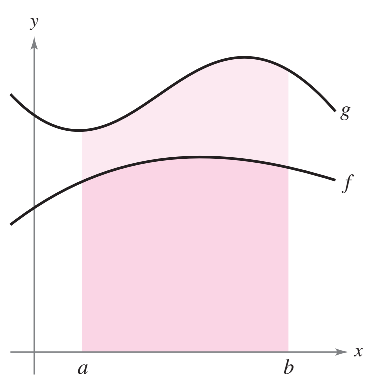
$$\begin{aligned}
\int_1^3 (-x^2 + 4x - 3)\,dx &= \int_1^3 (-x^2)\,dx + \int_1^3 4x\,dx + \int_1^3 (-3)\,dx \\
&= -\int_1^3 x^2\,dx + 4\int_1^3 x\,dx - 3\int_1^3 dx \\
&= -\left(\frac{26}{3}\right) + 4(4) - 3(2) \\
&= \frac{4}{3}
\end{aligned}$$

Try It | Exploration A | Exploration B

If f and g are continuous on the closed interval $[a, b]$ and

$0 \le f(x) \le g(x)$

for $a \le x \le b$, the following properties are true. First, the area of the region bounded by the graph of f and the x-axis (between a and b) must be nonnegative. Second, this area must be less than or equal to the area of the region bounded by the graph of g and the x-axis (between a and b), as shown in Figure 4.26. These two results are generalized in Theorem 4.8. (A proof of this theorem is given in Appendix A.)



$\displaystyle\int_a^b f(x)\,dx \le \int_a^b g(x)\,dx$

Figure 4.26

THEOREM 4.8 Preservation of Inequality

1. If f is integrable and nonnegative on the closed interval $[a, b]$, then

$$0 \leq \int_a^b f(x)\,dx.$$

2. If f and g are integrable on the closed interval $[a, b]$ and $f(x) \leq g(x)$ for every x in $[a, b]$, then

$$\int_a^b f(x)\,dx \leq \int_a^b g(x)\,dx.$$

Section 4.4 The Fundamental Theorem of Calculus

- **Evaluate a definite integral using the Fundamental Theorem of Calculus.**
- **Understand and use the Mean Value Theorem for Integrals.**
- **Find the average value of a function over a closed interval.**
- **Understand and use the Second Fundamental Theorem of Calculus.**

EXPLORATION

Integration and Antidifferentiation Throughout this chapter, you have been using the integral sign to denote an antiderivative (a family of functions) and a definite integral (a number).

Antidifferentiation: $\int f(x)\,dx$

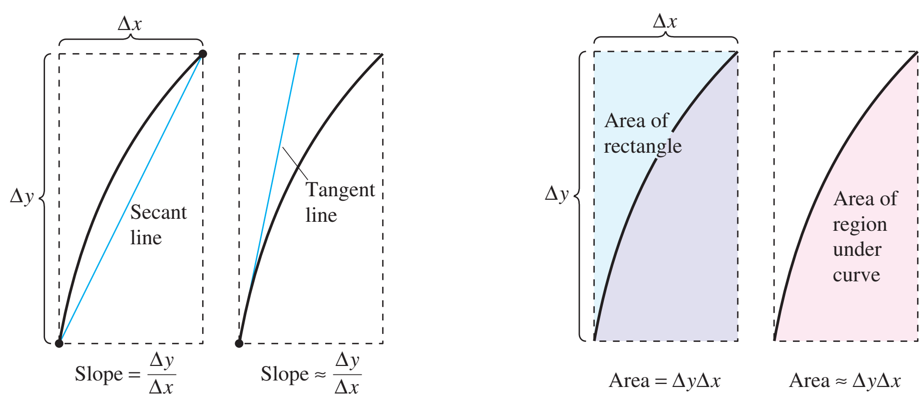
Definite integration: $\int_a^b f(x)\,dx$

The use of this same symbol for both operations makes it appear that they are related. In the early work with calculus, however, it was not known that the two operations were related. Do you think the symbol $\int$ was first applied to antidifferentiation or to definite integration? Explain your reasoning. (*Hint:* The symbol was first used by Leibniz and was derived from the letter S.)

The Fundamental Theorem of Calculus

You have now been introduced to the two major branches of calculus: differential calculus (introduced with the tangent line problem) and integral calculus (introduced with the area problem). At this point, these two problems might seem unrelated—but there is a very close connection. The connection was discovered independently by Isaac Newton and Gottfried Leibniz and is stated in a theorem that is appropriately called the **Fundamental Theorem of Calculus.**

Informally, the theorem states that differentiation and (definite) integration are inverse operations, in the same sense that division and multiplication are inverse operations. To see how Newton and Leibniz might have anticipated this relationship, consider the approximations shown in Figure 4.27. The slope of the tangent line was defined using the *quotient* $\Delta y/\Delta x$ (the slope of the secant line). Similarly, the area of a region under a curve was defined using the *product* $\Delta y \Delta x$ (the area of a rectangle). So, at least in the primitive approximation stage, the operations of differentiation and definite integration appear to have an inverse relationship in the same sense that division and multiplication are inverse operations. The Fundamental Theorem of Calculus states that the limit processes (used to define the derivative and definite integral) preserve this inverse relationship.



(a) Differentiation **(b)** Definite integration

Differentiation and definite integration have an "inverse" relationship.
Figure 4.27

THEOREM 4.9 The Fundamental Theorem of Calculus

If a function f is continuous on the closed interval $[a, b]$ and F is an antiderivative of f on the interval $[a, b]$, then

$$\int_a^b f(x)\,dx = F(b) - F(a).$$

Video

Proof The key to the proof is in writing the difference $F(b) - F(a)$ in a convenient form. Let Δ be the following partition of $[a, b]$.

$$a = x_0 < x_1 < x_2 < \cdot \cdot \cdot < x_{n-1} < x_n = b$$

By pairwise subtraction and addition of like terms, you can write

$$\begin{aligned} F(b) - F(a) &= F(x_n) - F(x_{n-1}) + F(x_{n-1}) - \cdot \cdot \cdot - F(x_1) + F(x_1) - F(x_0) \\ &= \sum_{i=1}^{n} [F(x_i) - F(x_{i-1})]. \end{aligned}$$

By the Mean Value Theorem, you know that there exists a number c_i in the ith subinterval such that

$$F'(c_i) = \frac{F(x_i) - F(x_{i-1})}{x_i - x_{i-1}}.$$

Because $F'(c_i) = f(c_i)$, you can let $\Delta x_i = x_i - x_{i-1}$ and obtain

$$F(b) - F(a) = \sum_{i=1}^{n} f(c_i)\Delta x_i.$$

This important equation tells you that by applying the Mean Value Theorem you can always find a collection of c_i's such that the *constant* $F(b) - F(a)$ is a Riemann sum of f on $[a, b]$. Taking the limit (as $\|\Delta\| \to 0$) produces

$$F(b) - F(a) = \int_a^b f(x)\, dx.$$

The following guidelines can help you understand the use of the Fundamental Theorem of Calculus.

Guidelines for Using the Fundamental Theorem of Calculus

1. *Provided you can find* an antiderivative of f, you now have a way to evaluate a definite integral without having to use the limit of a sum.
2. When applying the Fundamental Theorem of Calculus, the following notation is convenient.

$$\begin{aligned} \int_a^b f(x)\, dx &= F(x)\Big]_a^b \\ &= F(b) - F(a) \end{aligned}$$

For instance, to evaluate $\int_1^3 x^3\, dx$, you can write

$$\int_1^3 x^3\, dx = \frac{x^4}{4}\Big]_1^3 = \frac{3^4}{4} - \frac{1^4}{4} = \frac{81}{4} - \frac{1}{4} = 20.$$

3. It is not necessary to include a constant of integration C in the antiderivative because

$$\begin{aligned} \int_a^b f(x)\, dx &= \Big[F(x) + C\Big]_a^b \\ &= [F(b) + C] - [F(a) + C] \\ &= F(b) - F(a). \end{aligned}$$

EXAMPLE 1 Evaluating a Definite Integral

Evaluate each definite integral.

a. $\displaystyle\int_1^2 (x^2 - 3)\,dx$ **b.** $\displaystyle\int_1^4 3\sqrt{x}\,dx$ **c.** $\displaystyle\int_0^{\pi/4} \sec^2 x\,dx$

Solution

a. $\displaystyle\int_1^2 (x^2 - 3)\,dx = \left[\frac{x^3}{3} - 3x\right]_1^2 = \left(\frac{8}{3} - 6\right) - \left(\frac{1}{3} - 3\right) = -\frac{2}{3}$

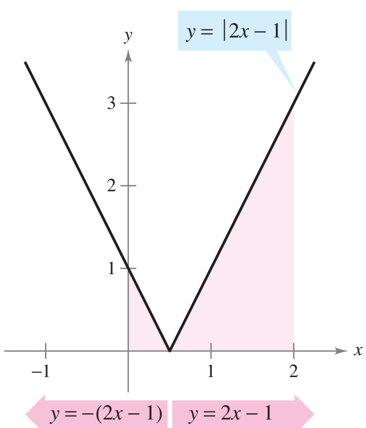
b. $\displaystyle\int_1^4 3\sqrt{x}\,dx = 3\int_1^4 x^{1/2}\,dx = 3\left[\frac{x^{3/2}}{3/2}\right]_1^4 = 2(4)^{3/2} - 2(1)^{3/2} = 14$

c. $\displaystyle\int_0^{\pi/4} \sec^2 x\,dx = \tan x\Big]_0^{\pi/4} = 1 - 0 = 1$

Try It | Exploration A | Open Exploration | Video

EXAMPLE 2 A Definite Integral Involving Absolute Value

Evaluate $\displaystyle\int_0^2 |2x - 1|\,dx.$



The definite integral of y on $[0, 2]$ is $\frac{5}{2}$.
Figure 4.28

Editable Graph

Solution Using Figure 4.28 and the definition of absolute value, you can rewrite the integrand as shown.

$$|2x - 1| = \begin{cases} -(2x - 1), & x < \frac{1}{2} \\ 2x - 1, & x \geq \frac{1}{2} \end{cases}$$

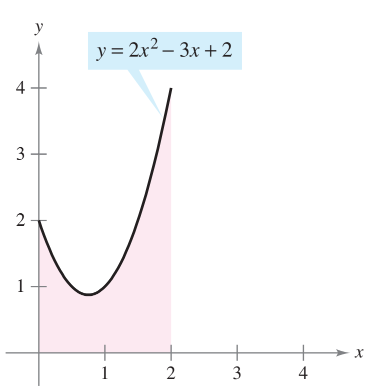
From this, you can rewrite the integral in two parts.

$$\begin{aligned} \int_0^2 |2x - 1|\,dx &= \int_0^{1/2} -(2x - 1)\,dx + \int_{1/2}^2 (2x - 1)\,dx \\ &= \Big[-x^2 + x\Big]_0^{1/2} + \Big[x^2 - x\Big]_{1/2}^2 \\ &= \left(-\frac{1}{4} + \frac{1}{2}\right) - (0 + 0) + (4 - 2) - \left(\frac{1}{4} - \frac{1}{2}\right) = \frac{5}{2} \end{aligned}$$

Try It | Exploration A

EXAMPLE 3 Using the Fundamental Theorem to Find Area

Find the area of the region bounded by the graph of $y = 2x^2 - 3x + 2$, the x-axis, and the vertical lines $x = 0$ and $x = 2$, as shown in Figure 4.29.



The area of the region bounded by the graph of y, the x-axis, $x = 0$, and $x = 2$ is $\frac{10}{3}$.
Figure 4.29

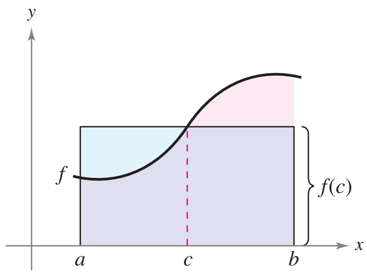
Editable Graph

Solution Note that $y > 0$ on the interval $[0, 2]$.

$$\begin{aligned} \text{Area} &= \int_0^2 (2x^2 - 3x + 2)\,dx && \text{Integrate between } x = 0 \text{ and } x = 2. \\ &= \left[\frac{2x^3}{3} - \frac{3x^2}{2} + 2x\right]_0^2 && \text{Find antiderivative.} \\ &= \left(\frac{16}{3} - 6 + 4\right) - (0 - 0 + 0) && \text{Apply Fundamental Theorem.} \\ &= \frac{10}{3} && \text{Simplify.} \end{aligned}$$

Try It | Exploration A

The Mean Value Theorem for Integrals



Mean value rectangle:

$$f(c)(b-a) = \int_a^b f(x)\,dx$$

Figure 4.30

In Section 4.2, you saw that the area of a region under a curve is greater than the area of an inscribed rectangle and less than the area of a circumscribed rectangle. The Mean Value Theorem for Integrals states that somewhere "between" the inscribed and circumscribed rectangles there is a rectangle whose area is precisely equal to the area of the region under the curve, as shown in Figure 4.30.

THEOREM 4.10 Mean Value Theorem for Integrals

If f is continuous on the closed interval $[a, b]$, then there exists a number c in the closed interval $[a, b]$ such that

$$\int_a^b f(x)\,dx = f(c)(b-a).$$

Video

Proof

Case 1: If f is constant on the interval $[a, b]$, the theorem is clearly valid because c can be any point in $[a, b]$.

Case 2: If f is not constant on $[a, b]$, then, by the Extreme Value Theorem, you can choose $f(m)$ and $f(M)$ to be the minimum and maximum values of f on $[a, b]$. Because $f(m) \le f(x) \le f(M)$ for all x in $[a, b]$, you can apply Theorem 4.8 to write the following.

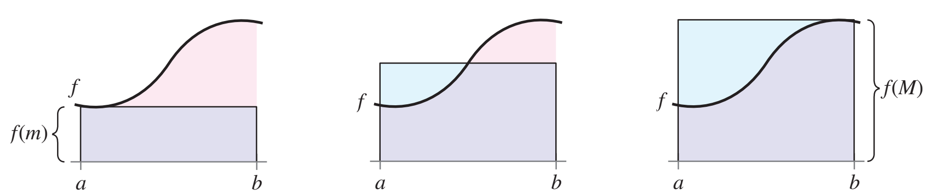
$$\int_a^b f(m)\,dx \le \int_a^b f(x)\,dx \le \int_a^b f(M)\,dx \qquad \text{See Figure 4.31.}$$

$$f(m)(b-a) \le \int_a^b f(x)\,dx \le f(M)(b-a)$$

$$f(m) \le \frac{1}{b-a}\int_a^b f(x)\,dx \le f(M)$$

From the third inequality, you can apply the Intermediate Value Theorem to conclude that there exists some c in $[a, b]$ such that

$$f(c) = \frac{1}{b-a}\int_a^b f(x)\,dx \quad \text{or} \quad f(c)(b-a) = \int_a^b f(x)\,dx.$$



Inscribed rectangle (less than actual area) $\int_a^b f(m)\,dx = f(m)(b-a)$

Mean value rectangle (equal to actual area) $\int_a^b f(x)\,dx$

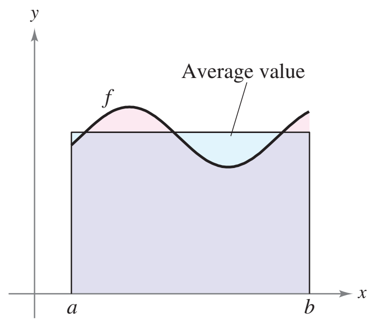
Circumscribed rectangle (greater than actual area) $\int_a^b f(M)\,dx = f(M)(b-a)$

Figure 4.31

NOTE Notice that Theorem 4.10 does not specify how to determine c. It merely guarantees the existence of at least one number c in the interval.

Average Value of a Function

The value of $f(c)$ given in the Mean Value Theorem for Integrals is called the **average value** of f on the interval $[a, b]$.



$$\text{Average value} = \frac{1}{b-a}\int_a^b f(x)\,dx$$

Figure 4.32

Definition of the Average Value of a Function on an Interval

If f is integrable on the closed interval $[a, b]$, then the **average value** of f on the interval is

$$\frac{1}{b-a}\int_a^b f(x)\,dx.$$

NOTE Notice in Figure 4.32 that the area of the region under the graph of f is equal to the area of the rectangle whose height is the average value.

To see why the average value of f is defined in this way, suppose that you partition $[a, b]$ into n subintervals of equal width $\Delta x = (b-a)/n$. If c_i is any point in the ith subinterval, the arithmetic average (or mean) of the function values at the c_i's is given by

$$a_n = \frac{1}{n}[f(c_1) + f(c_2) + \cdot \cdot \cdot + f(c_n)]. \qquad \text{Average of } f(c_1), \ldots, f(c_n)$$

By multiplying and dividing by $(b-a)$, you can write the average as

$$\begin{aligned} a_n &= \frac{1}{n}\sum_{i=1}^{n} f(c_i)\left(\frac{b-a}{b-a}\right) = \frac{1}{b-a}\sum_{i=1}^{n} f(c_i)\left(\frac{b-a}{n}\right) \\ &= \frac{1}{b-a}\sum_{i=1}^{n} f(c_i)\,\Delta x. \end{aligned}$$

Finally, taking the limit as $n \to \infty$ produces the average value of f on the interval $[a, b]$, as given in the definition above.

This development of the average value of a function on an interval is only one of many practical uses of definite integrals to represent summation processes. In Chapter 7, you will study other applications, such as volume, arc length, centers of mass, and work.

EXAMPLE 4 Finding the Average Value of a Function

Find the average value of $f(x) = 3x^2 - 2x$ on the interval $[1, 4]$.

Solution The average value is given by

$$\begin{aligned} \frac{1}{b-a}\int_a^b f(x)\,dx &= \frac{1}{3}\int_1^4 (3x^2 - 2x)\,dx \\ &= \frac{1}{3}\Big[x^3 - x^2\Big]_1^4 \\ &= \frac{1}{3}[64 - 16 - (1 - 1)] = \frac{48}{3} = 16. \end{aligned}$$

(See Figure 4.33.)

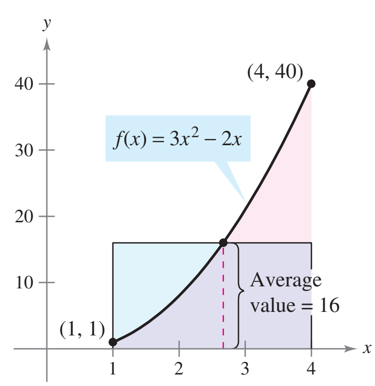


Figure 4.33

Editable Graph

Try It Exploration A Video

The first person to fly at a speed greater than the speed of sound was Charles Yeager. On October 14, 1947, flying in an *X-1* rocket plane at an altitude of 12.2 kilometers, Yeager was clocked at 295.9 meters per second. If Yeager had been flying at an altitude under 11.275 kilometers, his speed of 295.9 meters per second would not have "broken the sound barrier."

Video

EXAMPLE 5 The Speed of Sound

At different altitudes in Earth's atmosphere, sound travels at different speeds. The speed of sound $s(x)$ (in meters per second) can be modeled by

$$s(x) = \begin{cases} -4x + 341, & 0 \le x < 11.5 \\ 295, & 11.5 \le x < 22 \\ \frac{3}{4}x + 278.5, & 22 \le x < 32 \\ \frac{3}{2}x + 254.5, & 32 \le x < 50 \\ -\frac{3}{2}x + 404.5, & 50 \le x \le 80 \end{cases}$$

where x is the altitude in kilometers (see Figure 4.34). What is the average speed of sound over the interval $[0, 80]$?

Solution Begin by integrating $s(x)$ over the interval $[0, 80]$. To do this, you can break the integral into five parts.

$$\int_0^{11.5} s(x)\,dx = \int_0^{11.5} (-4x + 341)\,dx = \Big[-2x^2 + 341x\Big]_0^{11.5} = 3657$$

$$\int_{11.5}^{22} s(x)\,dx = \int_{11.5}^{22} (295)\,dx = \Big[295x\Big]_{11.5}^{22} = 3097.5$$

$$\int_{22}^{32} s(x)\,dx = \int_{22}^{32} \left(\tfrac{3}{4}x + 278.5\right)dx = \Big[\tfrac{3}{8}x^2 + 278.5x\Big]_{22}^{32} = 2987.5$$

$$\int_{32}^{50} s(x)\,dx = \int_{32}^{50} \left(\tfrac{3}{2}x + 254.5\right)dx = \Big[\tfrac{3}{4}x^2 + 254.5x\Big]_{32}^{50} = 5688$$

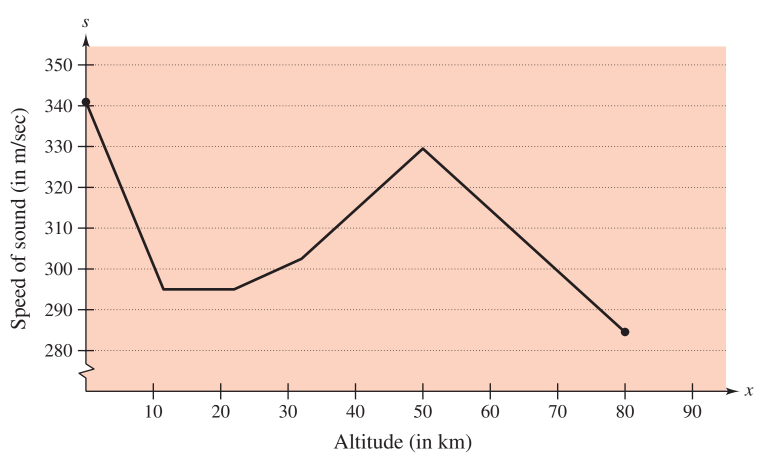
$$\int_{50}^{80} s(x)\,dx = \int_{50}^{80} \left(-\tfrac{3}{2}x + 404.5\right)dx = \Big[-\tfrac{3}{4}x^2 + 404.5x\Big]_{50}^{80} = 9210$$

By adding the values of the five integrals, you have

$$\int_0^{80} s(x)\,dx = 24{,}640.$$

So, the average speed of sound from an altitude of 0 kilometers to an altitude of 80 kilometers is

$$\text{Average speed} = \frac{1}{80}\int_0^{80} s(x)\,dx = \frac{24{,}640}{80} = 308 \text{ meters per second.}$$



Speed of sound depends on altitude.
Figure 4.34

Try It

Exploration A

The Second Fundamental Theorem of Calculus

Earlier you saw that the definite integral of f on the interval $[a, b]$ was defined using the constant b as the upper limit of integration and x as the variable of integration. However, a slightly different situation may arise in which the variable x is used as the upper limit of integration. To avoid the confusion of using x in two different ways, t is temporarily used as the variable of integration. (Remember that the definite integral is *not* a function of its variable of integration.)

The Definite Integral as a Number

Constant

$$\int_a^b f(x)\,dx$$

Constant

f is a function of x.

The Definite Integral as a Function of x

F is a function of x.

$$F(x) = \int_a^x f(t)\,dt$$

Constant

f is a function of t.

EXPLORATION

Use a graphing utility to graph the function

$$F(x) = \int_0^x \cos t\,dt$$

for $0 \le x \le \pi$. Do you recognize this graph? Explain.

EXAMPLE 6 The Definite Integral as a Function

Evaluate the function

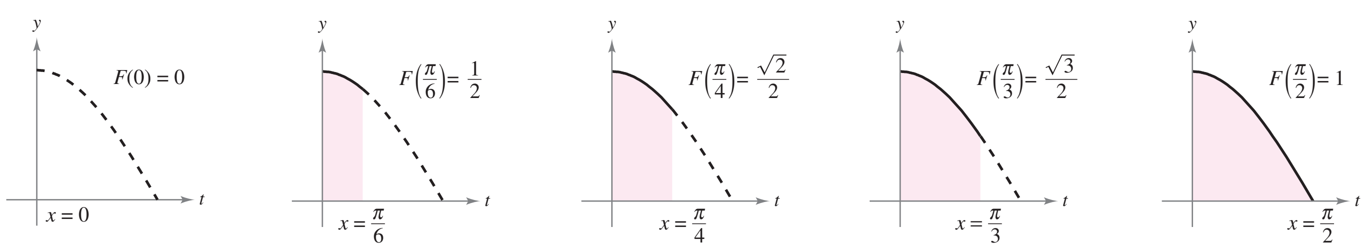
$$F(x) = \int_0^x \cos t\,dt$$

at $x = 0, \pi/6, \pi/4, \pi/3$, and $\pi/2$.

Solution You could evaluate five different definite integrals, one for each of the given upper limits. However, it is much simpler to fix x (as a constant) temporarily to obtain

$$\int_0^x \cos t\,dt = \sin t\Big]_0^x = \sin x - \sin 0 = \sin x.$$

Now, using $F(x) = \sin x$, you can obtain the results shown in Figure 4.35.



$F(x) = \int_0^x \cos t\,dt$ is the area under the curve $f(t) = \cos t$ from 0 to x.

Figure 4.35

Try It **Exploration A**

You can think of the function $F(x)$ as *accumulating* the area under the curve $f(t) = \cos t$ from $t = 0$ to $t = x$. For $x = 0$, the area is 0 and $F(0) = 0$. For $x = \pi/2$, $F(\pi/2) = 1$ gives the accumulated area under the cosine curve on the entire interval $[0, \pi/2]$. This interpretation of an integral as an **accumulation function** is used often in applications of integration.

In Example 6, note that the derivative of F is the original integrand (with only the variable changed). That is,

$$\frac{d}{dx}[F(x)] = \frac{d}{dx}[\sin x] = \frac{d}{dx}\left[\int_0^x \cos t\, dt\right] = \cos x.$$

This result is generalized in the following theorem, called the **Second Fundamental Theorem of Calculus.**

THEOREM 4.11 The Second Fundamental Theorem of Calculus

If f is continuous on an open interval I containing a, then, for every x in the interval,

$$\frac{d}{dx}\left[\int_a^x f(t)\, dt\right] = f(x).$$

Proof Begin by defining F as

$$F(x) = \int_a^x f(t)\, dt.$$

Then, by the definition of the derivative, you can write

$$\begin{aligned} F'(x) &= \lim_{\Delta x \to 0} \frac{F(x + \Delta x) - F(x)}{\Delta x} \\ &= \lim_{\Delta x \to 0} \frac{1}{\Delta x}\left[\int_a^{x+\Delta x} f(t)\, dt - \int_a^x f(t)\, dt\right] \\ &= \lim_{\Delta x \to 0} \frac{1}{\Delta x}\left[\int_a^{x+\Delta x} f(t)\, dt + \int_x^a f(t)\, dt\right] \\ &= \lim_{\Delta x \to 0} \frac{1}{\Delta x}\left[\int_x^{x+\Delta x} f(t)\, dt\right]. \end{aligned}$$

From the Mean Value Theorem for Integrals (assuming $\Delta x > 0$), you know there exists a number c in the interval $[x, x + \Delta x]$ such that the integral in the expression above is equal to $f(c)\,\Delta x$. Moreover, because $x \le c \le x + \Delta x$, it follows that $c \to x$ as $\Delta x \to 0$. So, you obtain

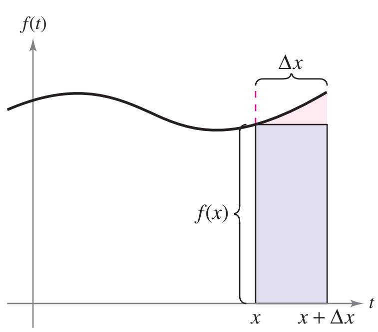
$$\begin{aligned} F'(x) &= \lim_{\Delta x \to 0} \left[\frac{1}{\Delta x} f(c)\,\Delta x\right] \\ &= \lim_{\Delta x \to 0} f(c) \\ &= f(x). \end{aligned}$$

A similar argument can be made for $\Delta x < 0$.

NOTE Using the area model for definite integrals, you can view the approximation

$$f(x)\,\Delta x \approx \int_x^{x+\Delta x} f(t)\, dt$$

as saying that the area of the rectangle of height $f(x)$ and width Δx is approximately equal to the area of the region lying between the graph of f and the x-axis on the interval $[x, x + \Delta x]$, as shown in Figure 4.36.



$f(x)\,\Delta x \approx \int_x^{x+\Delta x} f(t)\, dt$

Figure 4.36

Note that the Second Fundamental Theorem of Calculus tells you that if a function is continuous, you can be sure that it has an antiderivative. This antiderivative need not, however, be an elementary function. (Recall the discussion of elementary functions in Section P.3.)

EXAMPLE 7 Using the Second Fundamental Theorem of Calculus

Evaluate $\dfrac{d}{dx}\left[\displaystyle\int_0^x \sqrt{t^2+1}\,dt\right]$.

Solution Note that $f(t) = \sqrt{t^2+1}$ is continuous on the entire real line. So, using the Second Fundamental Theorem of Calculus, you can write

$$\frac{d}{dx}\left[\int_0^x \sqrt{t^2+1}\,dt\right] = \sqrt{x^2+1}.$$

The differentiation shown in Example 7 is a straightforward application of the Second Fundamental Theorem of Calculus. The next example shows how this theorem can be combined with the Chain Rule to find the derivative of a function.

Try It | Exploration A | Exploration B | Exploration C

Technology

EXAMPLE 8 Using the Second Fundamental Theorem of Calculus

Find the derivative of $F(x) = \displaystyle\int_{\pi/2}^{x^3} \cos t\,dt$.

Solution Using $u = x^3$, you can apply the Second Fundamental Theorem of Calculus with the Chain Rule as shown.

$$
\begin{aligned}
F'(x) &= \frac{dF}{du}\frac{du}{dx} && \text{Chain Rule} \\
&= \frac{d}{du}[F(x)]\frac{du}{dx} && \text{Definition of } \frac{dF}{du} \\
&= \frac{d}{du}\left[\int_{\pi/2}^{x^3} \cos t\,dt\right]\frac{du}{dx} && \text{Substitute } \int_{\pi/2}^{x^3} \cos t\,dt \text{ for } F(x). \\
&= \frac{d}{du}\left[\int_{\pi/2}^{u} \cos t\,dt\right]\frac{du}{dx} && \text{Substitute } u \text{ for } x^3. \\
&= (\cos u)(3x^2) && \text{Apply Second Fundamental Theorem of Calculus.} \\
&= (\cos x^3)(3x^2) && \text{Rewrite as function of } x.
\end{aligned}
$$

Try It | Exploration A

Because the integrand in Example 8 is easily integrated, you can verify the derivative as follows.

$$F(x) = \int_{\pi/2}^{x^3} \cos t\,dt = \sin t\Big]_{\pi/2}^{x^3} = \sin x^3 - \sin\frac{\pi}{2} = (\sin x^3) - 1$$

In this form, you can apply the Power Rule to verify that the derivative is the same as that obtained in Example 8.

$$F'(x) = (\cos x^3)(3x^2)$$

Section 4.5 Integration by Substitution

- **Use pattern recognition to find an indefinite integral.**
- **Use a change of variables to find an indefinite integral.**
- **Use the General Power Rule for Integration to find an indefinite integral.**
- **Use a change of variables to evaluate a definite integral.**
- **Evaluate a definite integral involving an even or odd function.**

Pattern Recognition

In this section you will study techniques for integrating composite functions. The discussion is split into two parts—*pattern recognition* and *change of variables*. Both techniques involve a ***u*-substitution.** With pattern recognition you perform the substitution mentally, and with change of variables you write the substitution steps.

The role of substitution in integration is comparable to the role of the Chain Rule in differentiation. Recall that for differentiable functions given by $y = F(u)$ and $u = g(x)$, the Chain Rule states that

$$\frac{d}{dx}[F(g(x))] = F'(g(x))g'(x).$$

From the definition of an antiderivative, it follows that

$$\begin{aligned}\int F'(g(x))g'(x)\,dx &= F(g(x)) + C\\ &= F(u) + C.\end{aligned}$$

These results are summarized in the following theorem.

NOTE The statement of Theorem 4.12 doesn't tell how to distinguish between $f(g(x))$ and $g'(x)$ in the integrand. As you become more experienced at integration, your skill in doing this will increase. Of course, part of the key is familiarity with derivatives.

THEOREM 4.12 Antidifferentiation of a Composite Function

Let g be a function whose range is an interval I, and let f be a function that is continuous on I. If g is differentiable on its domain and F is an antiderivative of f on I, then

$$\int f(g(x))g'(x)\,dx = F(g(x)) + C.$$

If $u = g(x)$, then $du = g'(x)\,dx$ and

$$\int f(u)\,du = F(u) + C.$$

STUDY TIP There are several techniques for applying substitution, each differing slightly from the others. However, you should remember that the goal is the same with every technique—*you are trying to find an antiderivative of the integrand.*

EXPLORATION

Recognizing Patterns The integrand in each of the following integrals fits the pattern $f(g(x))g'(x)$. Identify the pattern and use the result to evaluate the integral.

a. $\displaystyle\int 2x(x^2+1)^4\,dx$ **b.** $\displaystyle\int 3x^2\sqrt{x^3+1}\,dx$ **c.** $\displaystyle\int \sec^2 x(\tan x + 3)\,dx$

The next three integrals are similar to the first three. Show how you can multiply and divide by a constant to evaluate these integrals.

d. $\displaystyle\int x(x^2+1)^4\,dx$ **e.** $\displaystyle\int x^2\sqrt{x^3+1}\,dx$ **f.** $\displaystyle\int 2\sec^2 x(\tan x + 3)\,dx$

Examples 1 and 2 show how to apply Theorem 4.12 *directly*, by recognizing the presence of $f(g(x))$ and $g'(x)$. Note that the composite function in the integrand has an *outside function* f and an *inside function* g. Moreover, the derivative $g'(x)$ is present as a factor of the integrand.

Outside function

$$\int f(g(x))\underbrace{g'(x)}\,dx = F(g(x)) + C$$

Inside function

Derivative of inside function

EXAMPLE 1 Recognizing the $f(g(x))g'(x)$ Pattern

Find $\int (x^2 + 1)^2(2x)\,dx$.

Solution Letting $g(x) = x^2 + 1$, you obtain

$$g'(x) = 2x$$

and

$$f(g(x)) = f(x^2 + 1) = (x^2 + 1)^2.$$

From this, you can recognize that the integrand follows the $f(g(x))g'(x)$ pattern. Using the Power Rule for Integration and Theorem 4.12, you can write

$$\int \overbrace{(x^2 + 1)^2}^{f(g(x))}\overbrace{(2x)}^{g'(x)}\,dx = \frac{1}{3}(x^2 + 1)^3 + C.$$

Try using the Chain Rule to check that the derivative of $\frac{1}{3}(x^2 + 1)^3 + C$ is the integrand of the original integral.

TECHNOLOGY Try using a computer algebra system, such as *Maple*, *Derive*, *Mathematica*, *Mathcad*, or the *TI-89*, to solve the integrals given in Examples 1 and 2. Do you obtain the same antiderivatives that are listed in the examples?

Try It | Exploration A | Technology | Video

EXAMPLE 2 Recognizing the $f(g(x))g'(x)$ Pattern

Find $\int 5 \cos 5x\,dx$.

Solution Letting $g(x) = 5x$, you obtain

$$g'(x) = 5$$

and

$$f(g(x)) = f(5x) = \cos 5x.$$

From this, you can recognize that the integrand follows the $f(g(x))g'(x)$ pattern. Using the Cosine Rule for Integration and Theorem 4.12, you can write

$$\int \overbrace{(\cos 5x)}^{f(g(x))}\overbrace{(5)}^{g'(x)}\,dx = \sin 5x + C.$$

You can check this by differentiating $\sin 5x + C$ to obtain the original integrand.

Try It | Exploration A

The integrands in Examples 1 and 2 fit the $f(g(x))g'(x)$ pattern exactly—you only had to recognize the pattern. You can extend this technique considerably with the Constant Multiple Rule

$$\int kf(x)\,dx = k\int f(x)\,dx.$$

Many integrands contain the essential part (the variable part) of $g'(x)$ but are missing a constant multiple. In such cases, you can multiply and divide by the necessary constant multiple, as shown in Example 3.

EXAMPLE 3 Multiplying and Dividing by a Constant

Find $\int x(x^2+1)^2\,dx$.

Solution This is similar to the integral given in Example 1, except that the integrand is missing a factor of 2. Recognizing that $2x$ is the derivative of x^2+1, you can let $g(x) = x^2+1$ and supply the $2x$ as follows.

$$\begin{aligned}
\int x(x^2+1)^2\,dx &= \int (x^2+1)^2\left(\frac{1}{2}\right)(2x)\,dx && \text{Multiply and divide by 2.}\\
&= \frac{1}{2}\int \overbrace{(x^2+1)^2}^{f(g(x))}\,\overbrace{(2x)}^{g'(x)}\,dx && \text{Constant Multiple Rule}\\
&= \frac{1}{2}\left[\frac{(x^2+1)^3}{3}\right] + C && \text{Integrate.}\\
&= \frac{1}{6}(x^2+1)^3 + C && \text{Simplify.}
\end{aligned}$$

Try It | Exploration A | Exploration B

In practice, most people would not write as many steps as are shown in Example 3. For instance, you could evaluate the integral by simply writing

$$\begin{aligned}
\int x(x^2+1)^2\,dx &= \frac{1}{2}\int (x^2+1)^2\,2x\,dx\\
&= \frac{1}{2}\left[\frac{(x^2+1)^3}{3}\right] + C\\
&= \frac{1}{6}(x^2+1)^3 + C.
\end{aligned}$$

NOTE Be sure you see that the *Constant* Multiple Rule applies only to *constants*. You cannot multiply and divide by a variable and then move the variable outside the integral sign. For instance,

$$\int (x^2+1)^2\,dx \neq \frac{1}{2x}\int (x^2+1)^2(2x)\,dx.$$

After all, if it were legitimate to move variable quantities outside the integral sign, you could move the entire integrand out and simplify the whole process. But the result would be incorrect.

Change of Variables

With a formal **change of variables,** you completely rewrite the integral in terms of u and du (or any other convenient variable). Although this procedure can involve more written steps than the pattern recognition illustrated in Examples 1 to 3, it is useful for complicated integrands. The change of variable technique uses the Leibniz notation for the differential. That is, if $u = g(x)$, then $du = g'(x)\,dx$, and the integral in Theorem 4.12 takes the form

$$\int f(g(x))g'(x)\,dx = \int f(u)\,du = F(u) + C.$$

EXAMPLE 4 Change of Variables

Find $\int \sqrt{2x-1}\,dx$.

Solution First, let u be the inner function, $u = 2x - 1$. Then calculate the differential du to be $du = 2\,dx$. Now, using $\sqrt{2x-1} = \sqrt{u}$ and $dx = du/2$, substitute to obtain

$$\begin{aligned}
\int \sqrt{2x-1}\,dx &= \int \sqrt{u}\left(\frac{du}{2}\right) && \text{Integral in terms of } u\\
&= \frac{1}{2}\int u^{1/2}\,du && \text{Constant Multiple Rule}\\
&= \frac{1}{2}\left(\frac{u^{3/2}}{3/2}\right) + C && \text{Antiderivative in terms of } u\\
&= \frac{1}{3}u^{3/2} + C && \text{Simplify.}\\
&= \frac{1}{3}(2x-1)^{3/2} + C. && \text{Antiderivative in terms of } x
\end{aligned}$$

STUDY TIP Because integration is usually more difficult than differentiation, you should always check your answer to an integration problem by differentiating. For instance, in Example 4 you should differentiate $\frac{1}{3}(2x-1)^{3/2} + C$ to verify that you obtain the original integrand.

Try It | Exploration A

EXAMPLE 5 Change of Variables

Find $\int x\sqrt{2x-1}\,dx$.

Solution As in the previous example, let $u = 2x - 1$ and obtain $dx = du/2$. Because the integrand contains a factor of x, you must also solve for x in terms of u, as shown.

$$u = 2x - 1 \quad \Rightarrow \quad x = (u+1)/2 \qquad \text{Solve for } x \text{ in terms of } u.$$

Now, using substitution, you obtain

$$\begin{aligned}
\int x\sqrt{2x-1}\,dx &= \int \left(\frac{u+1}{2}\right)u^{1/2}\left(\frac{du}{2}\right)\\
&= \frac{1}{4}\int (u^{3/2} + u^{1/2})\,du\\
&= \frac{1}{4}\left(\frac{u^{5/2}}{5/2} + \frac{u^{3/2}}{3/2}\right) + C\\
&= \frac{1}{10}(2x-1)^{5/2} + \frac{1}{6}(2x-1)^{3/2} + C.
\end{aligned}$$

Try It | Exploration A | Exploration B | Open Exploration

To complete the change of variables in Example 5, you solved for x in terms of u. Sometimes this is very difficult. Fortunately it is not always necessary, as shown in the next example.

EXAMPLE 6 Change of Variables

Find $\int \sin^2 3x \cos 3x \, dx$.

Solution Because $\sin^2 3x = (\sin 3x)^2$, you can let $u = \sin 3x$. Then

$$du = (\cos 3x)(3)\, dx.$$

Now, because $\cos 3x \, dx$ is part of the original integral, you can write

$$\frac{du}{3} = \cos 3x \, dx.$$

STUDY TIP When making a change of variables, be sure that your answer is written using the same variables as in the original integrand. For instance, in Example 6, you should not leave your answer as

$$\tfrac{1}{9}u^3 + C$$

but rather, replace u by $\sin 3x$.

Substituting u and $du/3$ in the original integral yields

$$\begin{aligned}\int \sin^2 3x \cos 3x \, dx &= \int u^2 \frac{du}{3}\\ &= \frac{1}{3}\int u^2 \, du\\ &= \frac{1}{3}\left(\frac{u^3}{3}\right) + C\\ &= \frac{1}{9}\sin^3 3x + C.\end{aligned}$$

You can check this by differentiating.

$$\begin{aligned}\frac{d}{dx}\left[\frac{1}{9}\sin^3 3x\right] &= \left(\frac{1}{9}\right)(3)(\sin 3x)^2(\cos 3x)(3)\\ &= \sin^2 3x \cos 3x\end{aligned}$$

Because differentiation produces the original integrand, you know that you have obtained the correct antiderivative.

Try It | Exploration A

The steps used for integration by substitution are summarized in the following guidelines.

Guidelines for Making a Change of Variables

1. Choose a substitution $u = g(x)$. Usually, it is best to choose the *inner* part of a composite function, such as a quantity raised to a power.
2. Compute $du = g'(x)\, dx$.
3. Rewrite the integral in terms of the variable u.
4. Find the resulting integral in terms of u.
5. Replace u by $g(x)$ to obtain an antiderivative in terms of x.
6. Check your answer by differentiating.

The General Power Rule for Integration

One of the most common u-substitutions involves quantities in the integrand that are raised to a power. Because of the importance of this type of substitution, it is given a special name—the **General Power Rule for Integration.** A proof of this rule follows directly from the (simple) Power Rule for Integration, together with Theorem 4.12.

THEOREM 4.13 The General Power Rule for Integration

If g is a differentiable function of x, then

$$\int [g(x)]^n g'(x)\,dx = \frac{[g(x)]^{n+1}}{n+1} + C, \qquad n \neq -1.$$

Equivalently, if $u = g(x)$, then

$$\int u^n\,du = \frac{u^{n+1}}{n+1} + C, \qquad n \neq -1.$$

EXAMPLE 7 Substitution and the General Power Rule

a. $$\int 3(3x-1)^4\,dx = \int \overbrace{(3x-1)^4}^{u^4}\overbrace{(3)\,dx}^{du} = \overbrace{\frac{(3x-1)^5}{5}}^{u^5/5} + C$$

b. $$\int (2x+1)(x^2+x)\,dx = \int \overbrace{(x^2+x)^1}^{u^1}\overbrace{(2x+1)\,dx}^{du} = \overbrace{\frac{(x^2+x)^2}{2}}^{u^2/2} + C$$

c. $$\int 3x^2\sqrt{x^3-2}\,dx = \int \overbrace{(x^3-2)^{1/2}}^{u^{1/2}}\overbrace{(3x^2)\,dx}^{du} = \overbrace{\frac{(x^3-2)^{3/2}}{3/2}}^{u^{3/2}/(3/2)} + C = \frac{2}{3}(x^3-2)^{3/2} + C$$

d. $$\int \frac{-4x}{(1-2x^2)^2}\,dx = \int \overbrace{(1-2x^2)^{-2}}^{u^{-2}}\overbrace{(-4x)\,dx}^{du} = \overbrace{\frac{(1-2x^2)^{-1}}{-1}}^{u^{-1}/(-1)} + C = -\frac{1}{1-2x^2} + C$$

e. $$\int \cos^2 x \sin x\,dx = -\int \overbrace{(\cos x)^2}^{u^2}\overbrace{(-\sin x)\,dx}^{du} = -\overbrace{\frac{(\cos x)^3}{3}}^{u^3/3} + C$$

Try It | Exploration A | Exploration B

EXPLORATION

Suppose you were asked to find one of the following integrals. Which one would you choose? Explain your reasoning.

a. $\int \sqrt{x^3+1}\,dx$ or

$\int x^2\sqrt{x^3+1}\,dx$

b. $\int \tan(3x)\sec^2(3x)\,dx$ or

$\int \tan(3x)\,dx$

Some integrals whose integrands involve quantities raised to powers cannot be found by the General Power Rule. Consider the two integrals

$$\int x(x^2+1)^2\,dx \quad \text{and} \quad \int (x^2+1)^2\,dx.$$

The substitution $u = x^2 + 1$ works in the first integral but not in the second. In the second, the substitution fails because the integrand lacks the factor x needed for du. Fortunately, *for this particular integral*, you can expand the integrand as $(x^2+1)^2 = x^4 + 2x^2 + 1$ and use the (simple) Power Rule to integrate each term.

Change of Variables for Definite Integrals

When using u-substitution with a definite integral, it is often convenient to determine the limits of integration for the variable u rather than to convert the antiderivative back to the variable x and evaluate at the original limits. This change of variables is stated explicitly in the next theorem. The proof follows from Theorem 4.12 combined with the Fundamental Theorem of Calculus.

THEOREM 4.14 Change of Variables for Definite Integrals

If the function $u = g(x)$ has a continuous derivative on the closed interval $[a, b]$ and f is continuous on the range of g, then

$$\int_a^b f(g(x))g'(x)\,dx = \int_{g(a)}^{g(b)} f(u)\,du.$$

EXAMPLE 8 Change of Variables

Evaluate $\displaystyle\int_0^1 x(x^2+1)^3\,dx$.

Solution To evaluate this integral, let $u = x^2 + 1$. Then, you obtain

$$u = x^2 + 1 \implies du = 2x\,dx.$$

Before substituting, determine the new upper and lower limits of integration.

Lower Limit	*Upper Limit*
When $x = 0$, $u = 0^2 + 1 = 1$.	When $x = 1$, $u = 1^2 + 1 = 2$.

Now, you can substitute to obtain

$$\begin{aligned}\int_0^1 x(x^2+1)^3\,dx &= \frac{1}{2}\int_0^1 (x^2+1)^3(2x)\,dx && \text{Integration limits for } x\\ &= \frac{1}{2}\int_1^2 u^3\,du && \text{Integration limits for } u\\ &= \frac{1}{2}\left[\frac{u^4}{4}\right]_1^2\\ &= \frac{1}{2}\left(4 - \frac{1}{4}\right)\\ &= \frac{15}{8}.\end{aligned}$$

Try rewriting the antiderivative $\frac{1}{2}(u^4/4)$ in terms of the variable x and evaluate the definite integral at the original limits of integration, as shown.

$$\begin{aligned}\frac{1}{2}\left[\frac{u^4}{4}\right]_1^2 &= \frac{1}{2}\left[\frac{(x^2+1)^4}{4}\right]_0^1\\ &= \frac{1}{2}\left(4 - \frac{1}{4}\right) = \frac{15}{8}\end{aligned}$$

Notice that you obtain the same result.

Try It Exploration A Video

EXAMPLE 9 Change of Variables

Evaluate $A = \displaystyle\int_1^5 \frac{x}{\sqrt{2x-1}}\,dx.$

Solution To evaluate this integral, let $u = \sqrt{2x-1}$. Then, you obtain

$$\begin{aligned} u^2 &= 2x - 1 \\ u^2 + 1 &= 2x \\ \frac{u^2+1}{2} &= x \\ u\,du &= dx. \qquad \text{Differentiate each side.} \end{aligned}$$

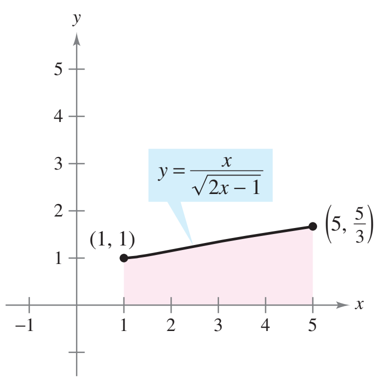
Before substituting, determine the new upper and lower limits of integration.

Lower Limit	*Upper Limit*
When $x = 1$, $u = \sqrt{2-1} = 1$.	When $x = 5$, $u = \sqrt{10-1} = 3$.

Now, substitute to obtain

$$\begin{aligned} \int_1^5 \frac{x}{\sqrt{2x-1}}\,dx &= \int_1^3 \frac{1}{u}\left(\frac{u^2+1}{2}\right) u\,du \\ &= \frac{1}{2}\int_1^3 (u^2+1)\,du \\ &= \frac{1}{2}\left[\frac{u^3}{3} + u\right]_1^3 \\ &= \frac{1}{2}\left(9 + 3 - \frac{1}{3} - 1\right) \\ &= \frac{16}{3}. \end{aligned}$$

Try It | Exploration A | Exploration B | Video



The region before substitution has an area of $\frac{16}{3}$.
Figure 4.37

Geometrically, you can interpret the equation

$$\int_1^5 \frac{x}{\sqrt{2x-1}}\,dx = \int_1^3 \frac{u^2+1}{2}\,du$$

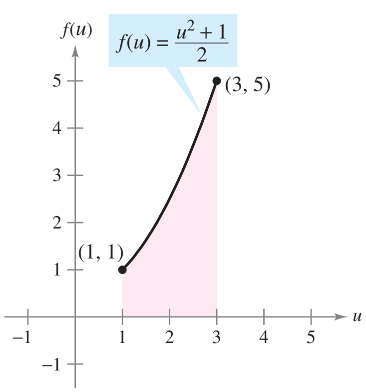
to mean that the two *different* regions shown in Figures 4.37 and 4.38 have the *same* area.

When evaluating definite integrals by substitution, it is possible for the upper limit of integration of the u-variable form to be smaller than the lower limit. If this happens, don't rearrange the limits. Simply evaluate as usual. For example, after substituting $u = \sqrt{1-x}$ in the integral

$$\int_0^1 x^2(1-x)^{1/2}\,dx$$

you obtain $u = \sqrt{1-1} = 0$ when $x = 1$, and $u = \sqrt{1-0} = 1$ when $x = 0$. So, the correct u-variable form of this integral is

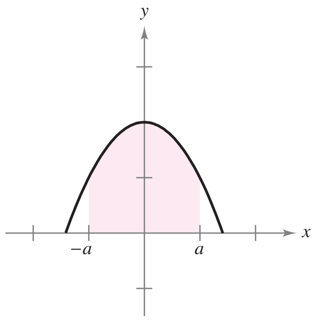
$$-2\int_1^0 (1-u^2)^2 u^2\,du.$$



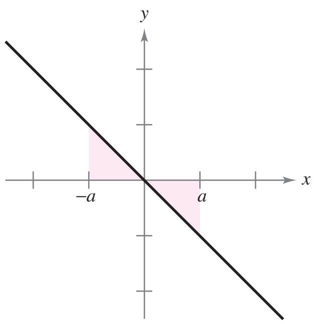
The region after substitution has an area of $\frac{16}{3}$.
Figure 4.38

Integration of Even and Odd Functions

Even with a change of variables, integration can be difficult. Occasionally, you can simplify the evaluation of a definite integral (over an interval that is symmetric about the y-axis or about the origin) by recognizing the integrand to be an even or odd function (see Figure 4.39).



Even function



Odd function
Figure 4.39

THEOREM 4.15 Integration of Even and Odd Functions

Let f be integrable on the closed interval $[-a, a]$.

1. If f is an *even* function, then $\displaystyle\int_{-a}^{a} f(x)\,dx = 2\int_{0}^{a} f(x)\,dx.$
2. If f is an *odd* function, then $\displaystyle\int_{-a}^{a} f(x)\,dx = 0.$

Proof Because f is even, you know that $f(x) = f(-x)$. Using Theorem 4.12 with the substitution $u = -x$ produces

$$\int_{-a}^{0} f(x)\,dx = \int_{a}^{0} f(-u)(-du) = -\int_{a}^{0} f(u)\,du = \int_{0}^{a} f(u)\,du = \int_{0}^{a} f(x)\,dx.$$

Finally, using Theorem 4.6, you obtain

$$\begin{aligned}\int_{-a}^{a} f(x)\,dx &= \int_{-a}^{0} f(x)\,dx + \int_{0}^{a} f(x)\,dx\\ &= \int_{0}^{a} f(x)\,dx + \int_{0}^{a} f(x)\,dx = 2\int_{0}^{a} f(x)\,dx.\end{aligned}$$

This proves the first property. The proof of the second property is left to you (see Exercise 133).

EXAMPLE 10 Integration of an Odd Function

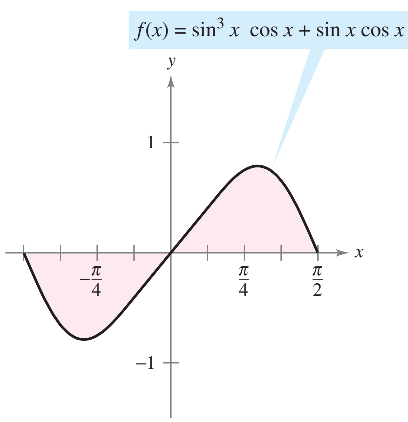
Evaluate $\displaystyle\int_{-\pi/2}^{\pi/2} (\sin^3 x \cos x + \sin x \cos x)\,dx.$

Solution Letting $f(x) = \sin^3 x \cos x + \sin x \cos x$ produces

$$\begin{aligned}f(-x) &= \sin^3(-x)\cos(-x) + \sin(-x)\cos(-x)\\ &= -\sin^3 x \cos x - \sin x \cos x = -f(x).\end{aligned}$$

So, f is an odd function, and because f is symmetric about the origin over $[-\pi/2, \pi/2]$, you can apply Theorem 4.15 to conclude that

$$\int_{-\pi/2}^{\pi/2} (\sin^3 x \cos x + \sin x \cos x)\,dx = 0.$$



Because f is an odd function,

$$\int_{-\pi/2}^{\pi/2} f(x)\,dx = 0.$$

Figure 4.40

Editable Graph

Try It

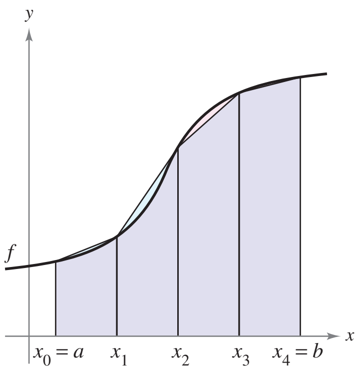
Exploration A

NOTE From Figure 4.40 you can see that the two regions on either side of the y-axis have the same area. However, because one lies below the x-axis and one lies above it, integration produces a cancellation effect. (More will be said about this in Section 7.1.)

Section 4.6 Numerical Integration

- **Approximate a definite integral using the Trapezoidal Rule.**
- **Approximate a definite integral using Simpson's Rule.**
- **Analyze the approximate errors in the Trapezoidal Rule and Simpson's Rule.**

The Trapezoidal Rule



The area of the region can be approximated using four trapezoids.
Figure 4.41

Some elementary functions simply do not have antiderivatives that are elementary functions. For example, there is no elementary function that has any of the following functions as its derivative.

$$\sqrt[3]{x}\sqrt{1-x}, \qquad \sqrt{x}\cos x, \qquad \frac{\cos x}{x}, \qquad \sqrt{1-x^3}, \qquad \sin x^2$$

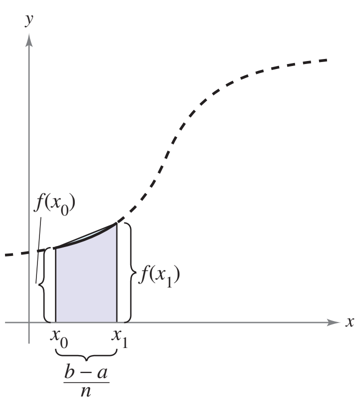
If you need to evaluate a definite integral involving a function whose antiderivative cannot be found, the Fundamental Theorem of Calculus cannot be applied, and you must resort to an approximation technique. Two such techniques are described in this section.

One way to approximate a definite integral is to use n trapezoids, as shown in Figure 4.41. In the development of this method, assume that f is continuous and positive on the interval $[a, b]$. So, the definite integral

$$\int_a^b f(x)\,dx$$

represents the area of the region bounded by the graph of f and the x-axis, from $x = a$ to $x = b$. First, partition the interval $[a, b]$ into n subintervals, each of width $\Delta x = (b - a)/n$, such that

$$a = x_0 < x_1 < x_2 < \cdot \cdot \cdot < x_n = b.$$



The area of the first trapezoid is
$\left[\dfrac{f(x_0) + f(x_1)}{2}\right]\left(\dfrac{b-a}{n}\right).$
Figure 4.42

Then form a trapezoid for each subinterval (see Figure 4.42). The area of the ith trapezoid is

$$\text{Area of } i\text{th trapezoid} = \left[\frac{f(x_{i-1}) + f(x_i)}{2}\right]\left(\frac{b-a}{n}\right).$$

This implies that the sum of the areas of the n trapezoids is

$$\begin{aligned}
\text{Area} &= \left(\frac{b-a}{n}\right)\left[\frac{f(x_0)+f(x_1)}{2} + \cdot \cdot \cdot + \frac{f(x_{n-1})+f(x_n)}{2}\right] \\
&= \left(\frac{b-a}{2n}\right)[f(x_0) + f(x_1) + f(x_1) + f(x_2) + \cdot \cdot \cdot + f(x_{n-1}) + f(x_n)] \\
&= \left(\frac{b-a}{2n}\right)[f(x_0) + 2f(x_1) + 2f(x_2) + \cdot \cdot \cdot + 2f(x_{n-1}) + f(x_n)].
\end{aligned}$$

Letting $\Delta x = (b - a)/n$, you can take the limit as $n \to \infty$ to obtain

$$\begin{aligned}
&\lim_{n\to\infty}\left(\frac{b-a}{2n}\right)[f(x_0) + 2f(x_1) + \cdot \cdot \cdot + 2f(x_{n-1}) + f(x_n)] \\
&\quad = \lim_{n\to\infty}\left[\frac{[f(a) - f(b)]\,\Delta x}{2} + \sum_{i=1}^{n} f(x_i)\,\Delta x\right] \\
&\quad = \lim_{n\to\infty}\frac{[f(a) - f(b)](b-a)}{2n} + \lim_{n\to\infty}\sum_{i=1}^{n} f(x_i)\,\Delta x \\
&\quad = 0 + \int_a^b f(x)\,dx.
\end{aligned}$$

The result is summarized in the following theorem.

THEOREM 4.16 The Trapezoidal Rule

Let f be continuous on $[a, b]$. The Trapezoidal Rule for approximating $\int_a^b f(x)\,dx$ is given by

$$\int_a^b f(x)\,dx \approx \frac{b-a}{2n}\left[f(x_0) + 2f(x_1) + 2f(x_2) + \cdot\cdot\cdot + 2f(x_{n-1}) + f(x_n)\right].$$

Moreover, as $n \to \infty$, the right-hand side approaches $\int_a^b f(x)\,dx$.

NOTE Observe that the coefficients in the Trapezoidal Rule have the following pattern.

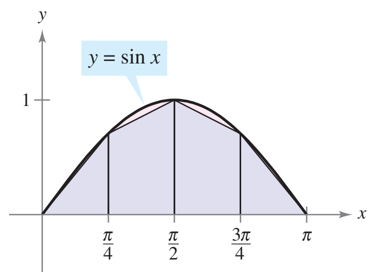
$$1 \quad 2 \quad 2 \quad 2 \quad . \; . \; . \quad 2 \quad 2 \quad 1$$

EXAMPLE 1 Approximation with the Trapezoidal Rule

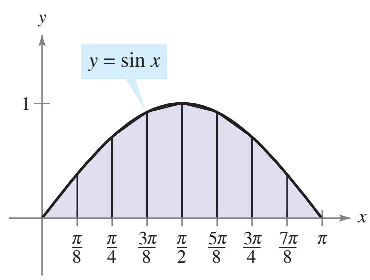
Use the Trapezoidal Rule to approximate

$$\int_0^\pi \sin x\,dx.$$

Compare the results for $n = 4$ and $n = 8$, as shown in Figure 4.43.



Four subintervals



Eight subintervals

Trapezoidal approximations
Figure 4.43

Solution When $n = 4$, $\Delta x = \pi/4$, and you obtain

$$\int_0^\pi \sin x\,dx \approx \frac{\pi}{8}\left(\sin 0 + 2\sin\frac{\pi}{4} + 2\sin\frac{\pi}{2} + 2\sin\frac{3\pi}{4} + \sin\pi\right)$$

$$= \frac{\pi}{8}\left(0 + \sqrt{2} + 2 + \sqrt{2} + 0\right) = \frac{\pi\left(1 + \sqrt{2}\right)}{4} \approx 1.896.$$

When $n = 8$, $\Delta x = \pi/8$, and you obtain

$$\int_0^\pi \sin x\,dx \approx \frac{\pi}{16}\left(\sin 0 + 2\sin\frac{\pi}{8} + 2\sin\frac{\pi}{4} + 2\sin\frac{3\pi}{8} + 2\sin\frac{\pi}{2} + 2\sin\frac{5\pi}{8} + 2\sin\frac{3\pi}{4} + 2\sin\frac{7\pi}{8} + \sin\pi\right)$$

$$= \frac{\pi}{16}\left(2 + 2\sqrt{2} + 4\sin\frac{\pi}{8} + 4\sin\frac{3\pi}{8}\right) \approx 1.974.$$

For this particular integral, you could have found an antiderivative and determined that the exact area of the region is 2.

Try It | Exploration A | Exploration B | Video

TECHNOLOGY Most graphing utilities and computer algebra systems have built-in programs that can be used to approximate the value of a definite integral. Try using such a program to approximate the integral in Example 1. How close is your approximation?

When you use such a program, you need to be aware of its limitations. Often, you are given no indication of the degree of accuracy of the approximation. Other times, you may be given an approximation that is completely wrong. For instance, try using a built-in numerical integration program to evaluate

$$\int_{-1}^2 \frac{1}{x}\,dx.$$

Your calculator should give an error message. Does yours?

It is interesting to compare the Trapezoidal Rule with the Midpoint Rule given in Section 4.2 (Exercises 63–66). For the Trapezoidal Rule, you average the function values at the endpoints of the subintervals, but for the Midpoint Rule you take the function values of the subinterval midpoints.

$$\int_a^b f(x)\,dx \approx \sum_{i=1}^{n} f\left(\frac{x_i + x_{i-1}}{2}\right)\Delta x \qquad \text{Midpoint Rule}$$

$$\int_a^b f(x)\,dx \approx \sum_{i=1}^{n} \left(\frac{f(x_i) + f(x_{i-1})}{2}\right)\Delta x \qquad \text{Trapezoidal Rule}$$

NOTE There are two important points that should be made concerning the Trapezoidal Rule (or the Midpoint Rule). First, the approximation tends to become more accurate as n increases. For instance, in Example 1, if $n = 16$, the Trapezoidal Rule yields an approximation of 1.994. Second, although you could have used the Fundamental Theorem to evaluate the integral in Example 1, this theorem cannot be used to evaluate an integral as simple as $\int_0^\pi \sin x^2\,dx$ because $\sin x^2$ has no elementary antiderivative. Yet, the Trapezoidal Rule can be applied easily to this integral.

Simpson's Rule

One way to view the trapezoidal approximation of a definite integral is to say that on each subinterval you approximate f by a *first*-degree polynomial. In Simpson's Rule, named after the English mathematician Thomas Simpson (1710–1761), you take this procedure one step further and approximate f by *second*-degree polynomials.

Before presenting Simpson's Rule, we list a theorem for evaluating integrals of polynomials of degree 2 (or less).

THEOREM 4.17 Integral of $p(x) = Ax^2 + Bx + C$

If $p(x) = Ax^2 + Bx + C$, then

$$\int_a^b p(x)\,dx = \left(\frac{b-a}{6}\right)\left[p(a) + 4p\left(\frac{a+b}{2}\right) + p(b)\right].$$

Proof

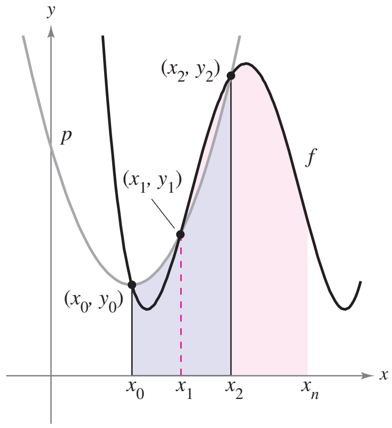
$$\begin{aligned}
\int_a^b p(x)\,dx &= \int_a^b (Ax^2 + Bx + C)\,dx \\
&= \left[\frac{Ax^3}{3} + \frac{Bx^2}{2} + Cx\right]_a^b \\
&= \frac{A(b^3 - a^3)}{3} + \frac{B(b^2 - a^2)}{2} + C(b - a) \\
&= \left(\frac{b-a}{6}\right)[2A(a^2 + ab + b^2) + 3B(b + a) + 6C]
\end{aligned}$$

By expansion and collection of terms, the expression inside the brackets becomes

$$\underbrace{(Aa^2 + Ba + C)}_{p(a)} + \underbrace{4\left[A\left(\frac{b+a}{2}\right)^2 + B\left(\frac{b+a}{2}\right) + C\right]}_{4p\left(\frac{a+b}{2}\right)} + \underbrace{(Ab^2 + Bb + C)}_{p(b)}$$

and you can write

$$\int_a^b p(x)\,dx = \left(\frac{b-a}{6}\right)\left[p(a) + 4p\left(\frac{a+b}{2}\right) + p(b)\right].$$



$\int_{x_0}^{x_2} p(x)\,dx \approx \int_{x_0}^{x_2} f(x)\,dx$

Figure 4.44

To develop Simpson's Rule for approximating a definite integral, you again partition the interval $[a, b]$ into n subintervals, each of width $\Delta x = (b - a)/n$. This time, however, n is required to be even, and the subintervals are grouped in pairs such that

$$a = \underbrace{x_0 < x_1 < x_2}_{[x_0, x_2]} < \underbrace{x_3 < x_4}_{[x_2, x_4]} < \cdots < \underbrace{x_{n-2} < x_{n-1} < x_n}_{[x_{n-2}, x_n]} = b.$$

On each (double) subinterval $[x_{i-2}, x_i]$, you can approximate f by a polynomial p of degree less than or equal to 2. (See Exercise 55.) For example, on the subinterval $[x_0, x_2]$, choose the polynomial of least degree passing through the points (x_0, y_0), (x_1, y_1), and (x_2, y_2), as shown in Figure 4.44. Now, using p as an approximation of f on this subinterval, you have, by Theorem 4.17,

$$\begin{aligned}\int_{x_0}^{x_2} f(x)\,dx \approx \int_{x_0}^{x_2} p(x)\,dx &= \frac{x_2 - x_0}{6}\left[p(x_0) + 4p\left(\frac{x_0 + x_2}{2}\right) + p(x_2)\right]\\ &= \frac{2[(b-a)/n]}{6}[p(x_0) + 4p(x_1) + p(x_2)]\\ &= \frac{b-a}{3n}[f(x_0) + 4f(x_1) + f(x_2)].\end{aligned}$$

Repeating this procedure on the entire interval $[a, b]$ produces the following theorem.

THEOREM 4.18 Simpson's Rule (n is even)

Let f be continuous on $[a, b]$. Simpson's Rule for approximating $\int_a^b f(x)\,dx$ is

$$\begin{aligned}\int_a^b f(x)\,dx \approx \frac{b-a}{3n}[f(x_0) &+ 4f(x_1) + 2f(x_2) + 4f(x_3) + \cdots\\ &+ 4f(x_{n-1}) + f(x_n)].\end{aligned}$$

Moreover, as $n \to \infty$, the right-hand side approaches $\int_a^b f(x)\,dx$.

NOTE Observe that the coefficients in Simpson's Rule have the following pattern.

1 4 2 4 2 4 . . . 4 2 4 1

In Example 1, the Trapezoidal Rule was used to estimate $\int_0^{\pi} \sin x\,dx$. In the next example, Simpson's Rule is applied to the same integral.

NOTE In Example 1, the Trapezoidal Rule with $n = 8$ approximated $\int_0^{\pi} \sin x\,dx$ as 1.974. In Example 2, Simpson's Rule with $n = 8$ gave an approximation of 2.0003. The antiderivative would produce the true value of 2.

EXAMPLE 2 Approximation with Simpson's Rule

Use Simpson's Rule to approximate

$$\int_0^{\pi} \sin x\,dx.$$

Compare the results for $n = 4$ and $n = 8$.

Solution When $n = 4$, you have

$$\int_0^{\pi} \sin x\,dx \approx \frac{\pi}{12}\left(\sin 0 + 4\sin\frac{\pi}{4} + 2\sin\frac{\pi}{2} + 4\sin\frac{3\pi}{4} + \sin\pi\right) \approx 2.005.$$

When $n = 8$, you have $\int_0^{\pi} \sin x\,dx \approx 2.0003$.

Try It Exploration A Open Exploration

Error Analysis

If you must use an approximation technique, it is important to know how accurate you can expect the approximation to be. The following theorem, which is listed without proof, gives the formulas for estimating the errors involved in the use of Simpson's Rule and the Trapezoidal Rule.

THEOREM 4.19 Errors in the Trapezoidal Rule and Simpson's Rule

If f has a continuous second derivative on $[a, b]$, then the error E in approximating $\int_a^b f(x)\,dx$ by the Trapezoidal Rule is

$$E \le \frac{(b-a)^3}{12n^2}\left[\max|f''(x)|\right], \quad a \le x \le b. \qquad \text{Trapezoidal Rule}$$

Moreover, if f has a continuous fourth derivative on $[a, b]$, then the error E in approximating $\int_a^b f(x)\,dx$ by Simpson's Rule is

$$E \le \frac{(b-a)^5}{180n^4}\left[\max|f^{(4)}(x)|\right], \quad a \le x \le b. \qquad \text{Simpson's Rule}$$

TECHNOLOGY If you have access to a computer algebra system, use it to evaluate the definite integral in Example 3. You should obtain a value of

$$\int_0^1 \sqrt{1+x^2}\,dx = \tfrac{1}{2}\left[\sqrt{2} + \ln\left(1+\sqrt{2}\right)\right]$$
$$\approx 1.14779.$$

("ln" represents the natural logarithmic function, which you will study in Section 5.1.)

Theorem 4.19 states that the errors generated by the Trapezoidal Rule and Simpson's Rule have upper bounds dependent on the extreme values of $f''(x)$ and $f^{(4)}(x)$ in the interval $[a, b]$. Furthermore, these errors can be made arbitrarily small by *increasing* n, provided that f'' and $f^{(4)}$ are continuous and therefore bounded in $[a, b]$.

EXAMPLE 3 The Approximate Error in the Trapezoidal Rule

Determine a value of n such that the Trapezoidal Rule will approximate the value of $\int_0^1 \sqrt{1+x^2}\,dx$ with an error that is less than 0.01.

Solution Begin by letting $f(x) = \sqrt{1+x^2}$ and finding the second derivative of f.

$$f'(x) = x(1+x^2)^{-1/2} \quad \text{and} \quad f''(x) = (1+x^2)^{-3/2}$$

The maximum value of $|f''(x)|$ on the interval $[0, 1]$ is $|f''(0)| = 1$. So, by Theorem 4.19, you can write

$$E \le \frac{(b-a)^3}{12n^2}|f''(0)| = \frac{1}{12n^2}(1) = \frac{1}{12n^2}.$$

To obtain an error E that is less than 0.01, you must choose n such that $1/(12n^2) \le 1/100$.

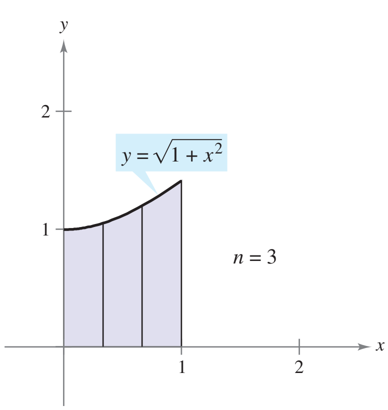
$$100 \le 12n^2 \quad \Longrightarrow \quad n \ge \sqrt{\frac{100}{12}} \approx 2.89$$

So, you can choose $n = 3$ (because n must be greater than or equal to 2.89) and apply the Trapezoidal Rule, as shown in Figure 4.45, to obtain

$$\int_0^1 \sqrt{1+x^2}\,dx \approx \frac{1}{6}\left[\sqrt{1+0^2} + 2\sqrt{1+\left(\tfrac{1}{3}\right)^2} + 2\sqrt{1+\left(\tfrac{2}{3}\right)^2} + \sqrt{1+1^2}\right]$$
$$\approx 1.154.$$

So, with an error no larger than 0.01, you know that

$$1.144 \le \int_0^1 \sqrt{1+x^2}\,dx \le 1.164.$$



$$1.144 \le \int_0^1 \sqrt{1+x^2}\,dx \le 1.164$$

Figure 4.45

Editable Graph

Try It Exploration A Exploration B

Section 5.1 The Natural Logarithmic Function: Differentiation

- **Develop and use properties of the natural logarithmic function.**
- **Understand the definition of the number *e*.**
- **Find derivatives of functions involving the natural logarithmic function.**

JOHN NAPIER (1550–1617)

Logarithms were invented by the Scottish mathematician John Napier. Although he did not introduce the *natural* logarithmic function, it is sometimes called the *Napierian* logarithm.

MathBio

The Natural Logarithmic Function

Recall that the General Power Rule

$$\int x^n\,dx = \frac{x^{n+1}}{n+1} + C, \quad n \neq -1 \qquad \text{General Power Rule}$$

has an important disclaimer—it doesn't apply when $n = -1$. Consequently, you have not yet found an antiderivative for the function $f(x) = 1/x$. In this section, you will use the Second Fundamental Theorem of Calculus to *define* such a function. This antiderivative is a function that you have not encountered previously in the text. It is neither algebraic nor trigonometric, but falls into a new class of functions called *logarithmic functions*. This particular function is the **natural logarithmic function.**

Definition of the Natural Logarithmic Function

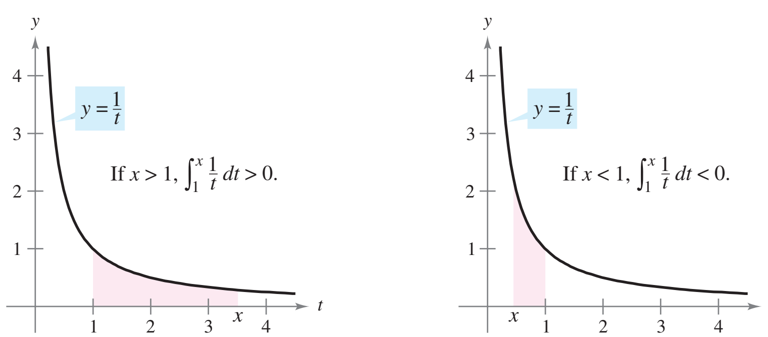
The **natural logarithmic function** is defined by

$$\ln x = \int_1^x \frac{1}{t}\,dt, \quad x > 0.$$

The domain of the natural logarithmic function is the set of all positive real numbers.

History Video

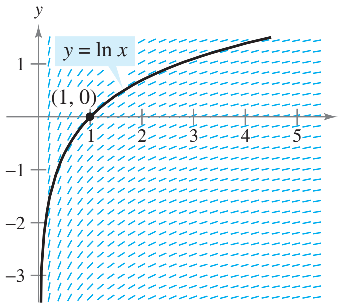
From the definition, you can see that $\ln x$ is positive for $x > 1$ and negative for $0 < x < 1$, as shown in Figure 5.1. Moreover, $\ln(1) = 0$, because the upper and lower limits of integration are equal when $x = 1$.



If $x > 1$, then $\ln x > 0$. If $0 < x < 1$, then $\ln x < 0$.
Figure 5.1

EXPLORATION

Graphing the Natural Logarithmic Function Using *only* the definition of the natural logarithmic function, sketch a graph of the function. Explain your reasoning.



Each small line segment has a slope of $\frac{1}{x}$.
Figure 5.2

NOTE Slope fields can be helpful in getting a visual perspective of the directions of the solutions of a differential equation.

To sketch the graph of $y = \ln x$, you can think of the natural logarithmic function as an *antiderivative* given by the differential equation

$$\frac{dy}{dx} = \frac{1}{x}.$$

Figure 5.2 is a computer-generated graph, called a *slope (or direction) field*, showing small line segments of slope $1/x$. The graph of $y = \ln x$ is the solution that passes through the point $(1, 0)$. You will study slope fields in Section 6.1.

The following theorem lists some basic properties of the natural logarithmic function.

THEOREM 5.1 Properties of the Natural Logarithmic Function

The natural logarithmic function has the following properties.

1. The domain is $(0, \infty)$ and the range is $(-\infty, \infty)$.
2. The function is continuous, increasing, and one-to-one.
3. The graph is concave downward.

Proof The domain of $f(x) = \ln x$ is $(0, \infty)$ by definition. Moreover, the function is continuous because it is differentiable. It is increasing because its derivative

$$f'(x) = \frac{1}{x} \qquad \text{First derivative}$$

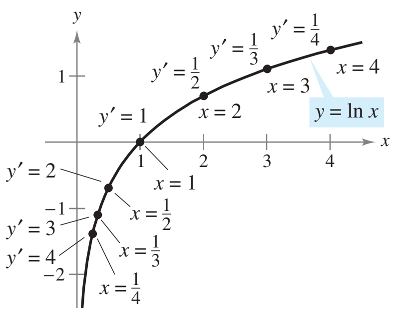
is positive for $x > 0$, as shown in Figure 5.3. It is concave downward because

$$f''(x) = -\frac{1}{x^2} \qquad \text{Second derivative}$$

is negative for $x > 0$. The proof that f is one-to-one is left as an exercise (see Exercise 111). The following limits imply that its range is the entire real line.

$$\lim_{x \to 0^+} \ln x = -\infty \quad \text{and} \quad \lim_{x \to \infty} \ln x = \infty$$

Verification of these two limits is given in Appendix A.



The natural logarithmic function is increasing, and its graph is concave downward.
Figure 5.3

Using the definition of the natural logarithmic function, you can prove several important properties involving operations with natural logarithms. If you are already familiar with logarithms, you will recognize that these properties are characteristic of all logarithms.

LOGARITHMS

Napier coined the term *logarithm*, from the two Greek words *logos* (or ratio) and *arithmos* (or number), to describe the theory that he spent 20 years developing and that first appeared in the book *Mirifici Logarithmorum canonis descriptio* (A Description of the Marvelous Rule of Logarithms).

THEOREM 5.2 Logarithmic Properties

If a and b are positive numbers and n is rational, then the following properties are true.

1. $\ln(1) = 0$
2. $\ln(ab) = \ln a + \ln b$
3. $\ln(a^n) = n \ln a$
4. $\ln\left(\frac{a}{b}\right) = \ln a - \ln b$

Proof The first property has already been discussed. The proof of the second property follows from the fact that two antiderivatives of the same function differ at most by a constant. From the Second Fundamental Theorem of Calculus and the definition of the natural logarithmic function, you know that

$$\frac{d}{dx}[\ln x] = \frac{d}{dx}\left[\int_1^x \frac{1}{t}\,dt\right] = \frac{1}{x}.$$

So, consider the two derivatives

$$\frac{d}{dx}[\ln(ax)] = \frac{a}{ax} = \frac{1}{x}$$

and

$$\frac{d}{dx}[\ln a + \ln x] = 0 + \frac{1}{x} = \frac{1}{x}.$$

Because $\ln(ax)$ and $(\ln a + \ln x)$ are both antiderivatives of $1/x$, they must differ at most by a constant.

$$\ln(ax) = \ln a + \ln x + C$$

By letting $x = 1$, you can see that $C = 0$. The third property can be proved similarly by comparing the derivatives of $\ln(x^n)$ and $n \ln x$. Finally, using the second and third properties, you can prove the fourth property.

$$\ln\left(\frac{a}{b}\right) = \ln[a(b^{-1})] = \ln a + \ln(b^{-1}) = \ln a - \ln b$$

Example 1 shows how logarithmic properties can be used to expand logarithmic expressions.

EXAMPLE 1 Expanding Logarithmic Expressions

a. $\ln\frac{10}{9} = \ln 10 - \ln 9$ — Property 4

b. $\ln\sqrt{3x+2} = \ln(3x+2)^{1/2}$ — Rewrite with rational exponent.

$= \frac{1}{2}\ln(3x+2)$ — Property 3

c. $\ln\frac{6x}{5} = \ln(6x) - \ln 5$ — Property 4

$= \ln 6 + \ln x - \ln 5$ — Property 2

d.
$$\ln\frac{(x^2+3)^2}{x\sqrt[3]{x^2+1}} = \ln(x^2+3)^2 - \ln\left(x\sqrt[3]{x^2+1}\right)$$
$$= 2\ln(x^2+3) - [\ln x + \ln(x^2+1)^{1/3}]$$
$$= 2\ln(x^2+3) - \ln x - \ln(x^2+1)^{1/3}$$
$$= 2\ln(x^2+3) - \ln x - \frac{1}{3}\ln(x^2+1)$$

Try It Exploration A

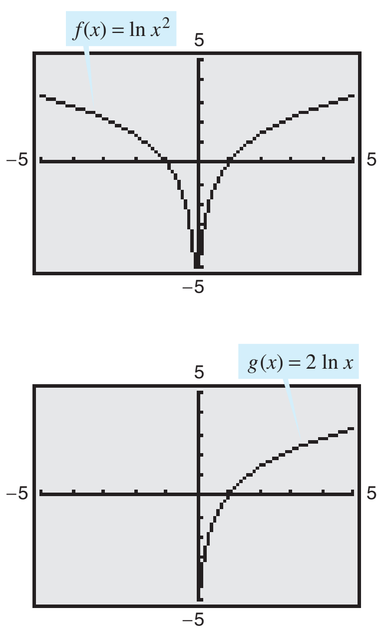


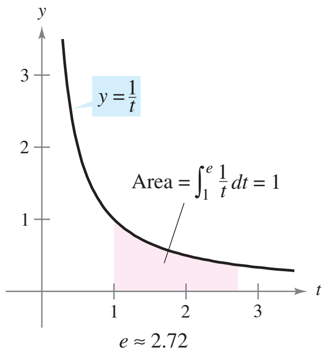
Figure 5.4

When using the properties of logarithms to rewrite logarithmic functions, you must check to see whether the domain of the rewritten function is the same as the domain of the original. For instance, the domain of $f(x) = \ln x^2$ is all real numbers except $x = 0$, and the domain of $g(x) = 2\ln x$ is all positive real numbers. (See Figure 5.4.)

The Number e

It is likely that you have studied logarithms in an algebra course. There, without the benefit of calculus, logarithms would have been defined in terms of a **base** number. For example, common logarithms have a base of 10 and therefore $\log_{10} 10 = 1$. (You will learn more about this in Section 5.5.)

The **base for the natural logarithm** is defined using the fact that the natural logarithmic function is continuous, is one-to-one, and has a range of $(-\infty, \infty)$. So, there must be a unique real number x such that $\ln x = 1$, as shown in Figure 5.5. This number is denoted by the letter e. It can be shown that e is irrational and has the following decimal approximation.



e is the base for the natural logarithm because $\ln e = 1$.
Figure 5.5

$$e \approx 2.71828182846$$

Definition of e

The letter e denotes the positive real number such that

$$\ln e = \int_1^e \frac{1}{t}\,dt = 1.$$

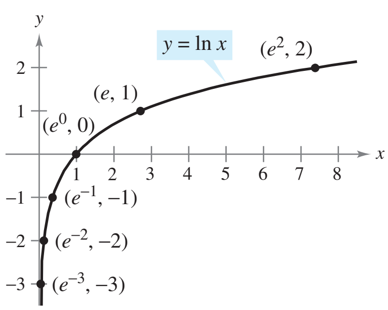
FOR FURTHER INFORMATION To learn more about the number e, see the article "Unexpected Occurrences of the Number e" by Harris S. Shultz and Bill Leonard in *Mathematics Magazine*.

MathArticle

Once you know that $\ln e = 1$, you can use logarithmic properties to evaluate the natural logarithms of several other numbers. For example, by using the property

$$\begin{aligned} \ln(e^n) &= n \ln e \\ &= n(1) \\ &= n \end{aligned}$$

you can evaluate $\ln(e^n)$ for various values of n, as shown in the table and in Figure 5.6.



If $x = e^n$, then $\ln x = n$.
Figure 5.6

Editable Graph

x	$\frac{1}{e^3} \approx 0.050$	$\frac{1}{e^2} \approx 0.135$	$\frac{1}{e} \approx 0.368$	$e^0 = 1$	$e \approx 2.718$	$e^2 \approx 7.389$
$\ln x$	-3	-2	-1	0	1	2

The logarithms shown in the table above are convenient because the x-values are integer powers of e. Most logarithmic expressions are, however, best evaluated with a calculator.

EXAMPLE 2 Evaluating Natural Logarithmic Expressions

a. $\ln 2 \approx 0.693$

b. $\ln 32 \approx 3.466$

c. $\ln 0.1 \approx -2.303$

Try It Exploration A

The Derivative of the Natural Logarithmic Function

The derivative of the natural logarithmic function is given in Theorem 5.3. The first part of the theorem follows from the definition of the natural logarithmic function as an antiderivative. The second part of the theorem is simply the Chain Rule version of the first part.

THEOREM 5.3 Derivative of the Natural Logarithmic Function

Let u be a differentiable function of x.

1. $\dfrac{d}{dx}[\ln x] = \dfrac{1}{x}, \quad x > 0$

2. $\dfrac{d}{dx}[\ln u] = \dfrac{1}{u}\dfrac{du}{dx} = \dfrac{u'}{u}, \quad u > 0$

Video

EXAMPLE 3 Differentiation of Logarithmic Functions

a. $\dfrac{d}{dx}[\ln(2x)] = \dfrac{u'}{u} = \dfrac{2}{2x} = \dfrac{1}{x}$ $\qquad u = 2x$

b. $\dfrac{d}{dx}[\ln(x^2+1)] = \dfrac{u'}{u} = \dfrac{2x}{x^2+1}$ $\qquad u = x^2 + 1$

c. $\dfrac{d}{dx}[x \ln x] = x\left(\dfrac{d}{dx}[\ln x]\right) + (\ln x)\left(\dfrac{d}{dx}[x]\right)$ $\qquad$ Product Rule

$$= x\left(\frac{1}{x}\right) + (\ln x)(1) = 1 + \ln x$$

d. $\dfrac{d}{dx}[(\ln x)^3] = 3(\ln x)^2 \dfrac{d}{dx}[\ln x]$ $\qquad$ Chain Rule

$$= 3(\ln x)^2 \frac{1}{x}$$

Try It Exploration A Exploration B

Open Exploration Video

EXPLORATION

Use a graphing utility to graph

$$y_1 = \frac{1}{x}$$

and

$$y_2 = \frac{d}{dx}[\ln x]$$

in the same viewing window, in which $0.1 \le x \le 5$ and $-2 \le y \le 8$. Explain why the graphs appear to be identical.

Napier used logarithmic properties to simplify *calculations* involving products, quotients, and powers. Of course, given the availability of calculators, there is now little need for this particular application of logarithms. However, there is great value in using logarithmic properties to simplify *differentiation* involving products, quotients, and powers.

EXAMPLE 4 Logarithmic Properties as Aids to Differentiation

Differentiate $f(x) = \ln\sqrt{x+1}$.

Solution Because

$$f(x) = \ln\sqrt{x+1} = \ln(x+1)^{1/2} = \frac{1}{2}\ln(x+1) \qquad \text{Rewrite before differentiating.}$$

you can write

$$f'(x) = \frac{1}{2}\left(\frac{1}{x+1}\right) = \frac{1}{2(x+1)}. \qquad \text{Differentiate.}$$

Try It Exploration A Video

EXAMPLE 5 Logarithmic Properties as Aids to Differentiation

Differentiate $f(x) = \ln \dfrac{x(x^2 + 1)^2}{\sqrt{2x^3 - 1}}$.

Solution

$$f(x) = \ln \frac{x(x^2 + 1)^2}{\sqrt{2x^3 - 1}} \qquad \text{Write original function.}$$

$$= \ln x + 2\ln(x^2 + 1) - \frac{1}{2}\ln(2x^3 - 1) \qquad \text{Rewrite before differentiating.}$$

$$f'(x) = \frac{1}{x} + 2\left(\frac{2x}{x^2 + 1}\right) - \frac{1}{2}\left(\frac{6x^2}{2x^3 - 1}\right) \qquad \text{Differentiate.}$$

$$= \frac{1}{x} + \frac{4x}{x^2 + 1} - \frac{3x^2}{2x^3 - 1} \qquad \text{Simplify.}$$

Try It

Exploration A

NOTE In Examples 4 and 5, be sure you see the benefit of applying logarithmic properties *before* differentiating. Consider, for instance, the difficulty of direct differentiation of the function given in Example 5.

On occasion, it is convenient to use logarithms as aids in differentiating *nonlogarithmic* functions. This procedure is called **logarithmic differentiation.**

EXAMPLE 6 Logarithmic Differentiation

Find the derivative of

$$y = \frac{(x - 2)^2}{\sqrt{x^2 + 1}}, \quad x \neq 2.$$

Solution Note that $y > 0$ for all $x \neq 2$. So, $\ln y$ is defined. Begin by taking the natural logarithm of each side of the equation. Then apply logarithmic properties and differentiate implicitly. Finally, solve for y'.

$$y = \frac{(x - 2)^2}{\sqrt{x^2 + 1}}, \quad x \neq 2 \qquad \text{Write original equation.}$$

$$\ln y = \ln \frac{(x - 2)^2}{\sqrt{x^2 + 1}} \qquad \text{Take natural log of each side.}$$

$$\ln y = 2\ln(x - 2) - \frac{1}{2}\ln(x^2 + 1) \qquad \text{Logarithmic properties}$$

$$\frac{y'}{y} = 2\left(\frac{1}{x - 2}\right) - \frac{1}{2}\left(\frac{2x}{x^2 + 1}\right) \qquad \text{Differentiate.}$$

$$= \frac{2}{x - 2} - \frac{x}{x^2 + 1} \qquad \text{Simplify.}$$

$$y' = y\left(\frac{2}{x - 2} - \frac{x}{x^2 + 1}\right) \qquad \text{Solve for } y'.$$

$$= \frac{(x - 2)^2}{\sqrt{x^2 + 1}}\left[\frac{x^2 + 2x + 2}{(x - 2)(x^2 + 1)}\right] \qquad \text{Substitute for } y.$$

$$= \frac{(x - 2)(x^2 + 2x + 2)}{(x^2 + 1)^{3/2}} \qquad \text{Simplify.}$$

Try It

Exploration A

The editable graph feature below allows you to edit the graph of a function.

Editable Graph

Because the natural logarithm is undefined for negative numbers, you will often encounter expressions of the form $\ln|u|$. The following theorem states that you can differentiate functions of the form $y = \ln|u|$ as if the absolute value sign were not present.

THEOREM 5.4 Derivative Involving Absolute Value

If u is a differentiable function of x such that $u \neq 0$, then

$$\frac{d}{dx}[\ln|u|] = \frac{u'}{u}.$$

Proof If $u > 0$, then $|u| = u$, and the result follows from Theorem 5.3. If $u < 0$, then $|u| = -u$, and you have

$$\begin{aligned}\frac{d}{dx}[\ln|u|] &= \frac{d}{dx}[\ln(-u)]\\ &= \frac{-u'}{-u}\\ &= \frac{u'}{u}.\end{aligned}$$

EXAMPLE 7 Derivative Involving Absolute Value

Find the derivative of

$$f(x) = \ln|\cos x|.$$

Solution Using Theorem 5.4, let $u = \cos x$ and write

$$\begin{aligned}\frac{d}{dx}[\ln|\cos x|] &= \frac{u'}{u} && \frac{d}{dx}[\ln|u|] = \frac{u'}{u}\\ &= \frac{-\sin x}{\cos x} && u = \cos x\\ &= -\tan x. && \text{Simplify.}\end{aligned}$$

Try It   Exploration A

The editable graph feature below allows you to edit the graph of a function.

Editable Graph

EXAMPLE 8 Finding Relative Extrema

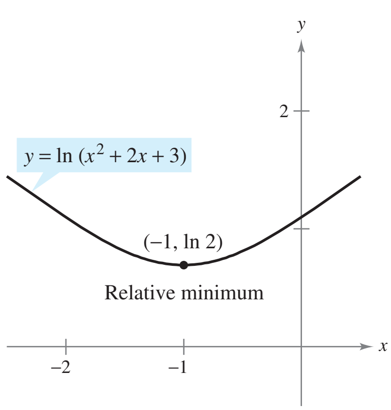
Locate the relative extrema of

$$y = \ln(x^2 + 2x + 3).$$

Solution Differentiating y, you obtain

$$\frac{dy}{dx} = \frac{2x + 2}{x^2 + 2x + 3}.$$

Because $dy/dx = 0$ when $x = -1$, you can apply the First Derivative Test and conclude that the point $(-1, \ln 2)$ is a relative minimum. Because there are no other critical points, it follows that this is the only relative extremum (see Figure 5.7).



The derivative of y changes from negative to positive at $x = -1$.
Figure 5.7

Editable Graph

Try It   Exploration A

Section 5.2 The Natural Logarithmic Function: Integration

- **Use the Log Rule for Integration to integrate a rational function.**
- **Integrate trigonometric functions.**

EXPLORATION

Integrating Rational Functions
Early in Chapter 4, you learned rules that allowed you to integrate *any* polynomial function. The Log Rule presented in this section goes a long way toward enabling you to integrate rational functions. For instance, each of the following functions can be integrated with the Log Rule.

$\dfrac{2}{x}$	Example 1
$\dfrac{1}{4x-1}$	Example 2
$\dfrac{x}{x^2+1}$	Example 3
$\dfrac{3x^2+1}{x^3+x}$	Example 4(a)
$\dfrac{x+1}{x^2+2x}$	Example 4(c)
$\dfrac{1}{3x+2}$	Example 4(d)
$\dfrac{x^2+x+1}{x^2+1}$	Example 5
$\dfrac{2x}{(x+1)^2}$	Example 6

There are still some rational functions that cannot be integrated using the Log Rule. Give examples of these functions, and explain your reasoning.

Log Rule for Integration

The differentiation rules

$$\frac{d}{dx}[\ln|x|] = \frac{1}{x} \qquad \text{and} \qquad \frac{d}{dx}[\ln|u|] = \frac{u'}{u}$$

that you studied in the preceding section produce the following integration rule.

THEOREM 5.5 Log Rule for Integration

Let u be a differentiable function of x.

1. $\displaystyle\int \frac{1}{x}\,dx = \ln|x| + C$ **2.** $\displaystyle\int \frac{1}{u}\,du = \ln|u| + C$

Video Video

Because $du = u'\,dx$, the second formula can also be written as

$$\int \frac{u'}{u}\,dx = \ln|u| + C. \qquad \text{Alternative form of Log Rule}$$

EXAMPLE 1 Using the Log Rule for Integration

$$\begin{aligned}\int \frac{2}{x}\,dx &= 2\int \frac{1}{x}\,dx && \text{Constant Multiple Rule}\\ &= 2\ln|x| + C && \text{Log Rule for Integration}\\ &= \ln(x^2) + C && \text{Property of logarithms}\end{aligned}$$

Because x^2 cannot be negative, the absolute value is unnecessary in the final form of the antiderivative.

Try It Exploration A

EXAMPLE 2 Using the Log Rule with a Change of Variables

Find $\displaystyle\int \frac{1}{4x-1}\,dx.$

Solution If you let $u = 4x - 1$, then $du = 4\,dx$.

$$\begin{aligned}\int \frac{1}{4x-1}\,dx &= \frac{1}{4}\int\left(\frac{1}{4x-1}\right)4\,dx && \text{Multiply and divide by 4.}\\ &= \frac{1}{4}\int \frac{1}{u}\,du && \text{Substitute: } u = 4x-1.\\ &= \frac{1}{4}\ln|u| + C && \text{Apply Log Rule.}\\ &= \frac{1}{4}\ln|4x-1| + C && \text{Back-substitute.}\end{aligned}$$

Try It Exploration A Video

Example 3 uses the alternative form of the Log Rule. To apply this rule, look for quotients in which the numerator is the derivative of the denominator.

EXAMPLE 3 Finding Area with the Log Rule

Find the area of the region bounded by the graph of

$$y = \frac{x}{x^2 + 1}$$

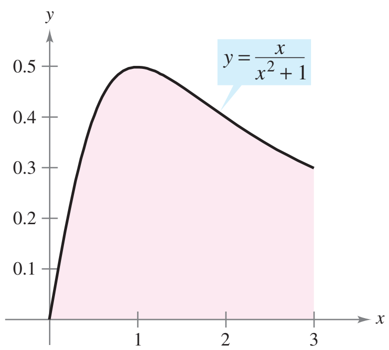
the x-axis, and the line $x = 3$.

Solution From Figure 5.8, you can see that the area of the region is given by the definite integral

$$\int_0^3 \frac{x}{x^2 + 1}\, dx.$$

If you let $u = x^2 + 1$, then $u' = 2x$. To apply the Log Rule, multiply and divide by 2 as shown.

$$\begin{aligned} \int_0^3 \frac{x}{x^2+1}\,dx &= \frac{1}{2}\int_0^3 \frac{2x}{x^2+1}\,dx && \text{Multiply and divide by 2.} \\ &= \frac{1}{2}\Big[\ln(x^2+1)\Big]_0^3 && \int \frac{u'}{u}\,dx = \ln|u| + C \\ &= \frac{1}{2}(\ln 10 - \ln 1) \\ &= \frac{1}{2}\ln 10 && \ln 1 = 0 \\ &\approx 1.151 \end{aligned}$$



Area $= \int_0^3 \frac{x}{x^2 + 1}\, dx$

The area of the region bounded by the graph of y, the x-axis, and $x = 3$ is $\frac{1}{2}\ln 10$.

Figure 5.8

Editable Graph

Try It Exploration A

EXAMPLE 4 Recognizing Quotient Forms of the Log Rule

a. $\displaystyle\int \frac{3x^2 + 1}{x^3 + x}\, dx = \ln|x^3 + x| + C$ $\qquad u = x^3 + x$

b. $\displaystyle\int \frac{\sec^2 x}{\tan x}\, dx = \ln|\tan x| + C$ $\qquad u = \tan x$

c. $\displaystyle\int \frac{x + 1}{x^2 + 2x}\, dx = \frac{1}{2}\int \frac{2x + 2}{x^2 + 2x}\, dx$ $\qquad u = x^2 + 2x$

$$= \frac{1}{2}\ln|x^2 + 2x| + C$$

d. $\displaystyle\int \frac{1}{3x + 2}\, dx = \frac{1}{3}\int \frac{3}{3x + 2}\, dx$ $\qquad u = 3x + 2$

$$= \frac{1}{3}\ln|3x + 2| + C$$

Try It Exploration A Exploration B

With antiderivatives involving logarithms, it is easy to obtain forms that look quite different but are still equivalent. For instance, which of the following are equivalent to the antiderivative listed in Example 4(d)?

$$\ln|(3x + 2)^{1/3}| + C, \quad \frac{1}{3}\ln\left|x + \tfrac{2}{3}\right| + C, \quad \ln|3x + 2|^{1/3} + C$$

Integrals to which the Log Rule can be applied often appear in disguised form. For instance, if a rational function has a *numerator of degree greater than or equal to that of the denominator*, division may reveal a form to which you can apply the Log Rule. This is shown in Example 5.

EXAMPLE 5 Using Long Division Before Integrating

Find $\displaystyle\int \frac{x^2 + x + 1}{x^2 + 1}\,dx.$

Solution Begin by using long division to rewrite the integrand.

$$\frac{x^2 + x + 1}{x^2 + 1} \quad \Rightarrow \quad x^2 + 1 \,\overline{)\, x^2 + x + 1} \;\; \begin{array}{r} 1 \\ \underline{x^2 \qquad + 1} \\ x \end{array} \quad \Rightarrow \quad 1 + \frac{x}{x^2 + 1}$$

Now, you can integrate to obtain

$$\begin{aligned}
\int \frac{x^2 + x + 1}{x^2 + 1}\,dx &= \int \left(1 + \frac{x}{x^2 + 1}\right) dx && \text{Rewrite using long division.} \\
&= \int dx + \frac{1}{2}\int \frac{2x}{x^2 + 1}\,dx && \text{Rewrite as two integrals.} \\
&= x + \frac{1}{2}\ln(x^2 + 1) + C. && \text{Integrate.}
\end{aligned}$$

Check this result by differentiating to obtain the original integrand.

Try It | Exploration A | Open Exploration

The next example gives another instance in which the use of the Log Rule is disguised. In this case, a change of variables helps you recognize the Log Rule.

EXAMPLE 6 Change of Variables with the Log Rule

Find $\displaystyle\int \frac{2x}{(x + 1)^2}\,dx.$

Solution If you let $u = x + 1$, then $du = dx$ and $x = u - 1$.

$$\begin{aligned}
\int \frac{2x}{(x + 1)^2}\,dx &= \int \frac{2(u - 1)}{u^2}\,du && \text{Substitute.} \\
&= 2\int \left(\frac{u}{u^2} - \frac{1}{u^2}\right) du && \text{Rewrite as two fractions.} \\
&= 2\int \frac{du}{u} - 2\int u^{-2}\,du && \text{Rewrite as two integrals.} \\
&= 2\ln|u| - 2\left(\frac{u^{-1}}{-1}\right) + C && \text{Integrate.} \\
&= 2\ln|u| + \frac{2}{u} + C && \text{Simplify.} \\
&= 2\ln|x + 1| + \frac{2}{x + 1} + C && \text{Back-substitute.}
\end{aligned}$$

Check this result by differentiating to obtain the original integrand.

Try It | Exploration A | Exploration B

TECHNOLOGY If you have access to a computer algebra system, use it to find the indefinite integrals in Examples 5 and 6. How does the form of the antiderivative that it gives you compare with that given in Examples 5 and 6?

As you study the methods shown in Examples 5 and 6, be aware that both methods involve rewriting a disguised integrand so that it fits one or more of the basic integration formulas. Throughout the remaining sections of Chapter 5 and in Chapter 8, much time will be devoted to integration techniques. To master these techniques, you must recognize the "form-fitting" nature of integration. In this sense, integration is not nearly as straightforward as differentiation. Differentiation takes the form

"Here is the question; what is the answer?"

Integration is more like

"Here is the answer; what is the question?"

The following are guidelines you can use for integration.

STUDY TIP Keep in mind that you can check your answer to an integration problem by differentiating the answer. For instance, in Example 7, the derivative of $y = \ln|\ln x| + C$ is $y' = 1/(x \ln x)$.

Guidelines for Integration

1. Learn a basic list of integration formulas. (Including those given in this section, you now have 12 formulas: the Power Rule, the Log Rule, and ten trigonometric rules. By the end of Section 5.7, this list will have expanded to 20 basic rules.)
2. Find an integration formula that resembles all or part of the integrand, and, by trial and error, find a choice of u that will make the integrand conform to the formula.
3. If you cannot find a u-substitution that works, try altering the integrand. You might try a trigonometric identity, multiplication and division by the same quantity, or addition and subtraction of the same quantity. Be creative.
4. If you have access to computer software that will find antiderivatives symbolically, use it.

EXAMPLE 7 *u*-Substitution and the Log Rule

Solve the differential equation $\dfrac{dy}{dx} = \dfrac{1}{x \ln x}$.

Solution The solution can be written as an indefinite integral.

$$y = \int \frac{1}{x \ln x}\, dx$$

Because the integrand is a quotient whose denominator is raised to the first power, you should try the Log Rule. There are three basic choices for u. The choices $u = x$ and $u = x \ln x$ fail to fit the u'/u form of the Log Rule. However, the third choice does fit. Letting $u = \ln x$ produces $u' = 1/x$, and you obtain the following.

$$\int \frac{1}{x \ln x}\, dx = \int \frac{1/x}{\ln x}\, dx \qquad \text{Divide numerator and denominator by } x.$$

$$= \int \frac{u'}{u}\, dx \qquad \text{Substitute: } u = \ln x.$$

$$= \ln|u| + C \qquad \text{Apply Log Rule.}$$

$$= \ln|\ln x| + C \qquad \text{Back-substitute.}$$

So, the solution is $y = \ln|\ln x| + C$.

Try It Exploration A

Integrals of Trigonometric Functions

In Section 4.1, you looked at six trigonometric integration rules—the six that correspond directly to differentiation rules. With the Log Rule, you can now complete the set of basic trigonometric integration formulas.

EXAMPLE 8 Using a Trigonometric Identity

Find $\int \tan x \, dx$.

Solution This integral does not seem to fit any formulas on our basic list. However, by using a trigonometric identity, you obtain

$$\int \tan x \, dx = \int \frac{\sin x}{\cos x} \, dx.$$

Knowing that $D_x[\cos x] = -\sin x$, you can let $u = \cos x$ and write

$$\begin{aligned}
\int \tan x \, dx &= -\int \frac{-\sin x}{\cos x} \, dx && \text{Trigonometric identity} \\
&= -\int \frac{u'}{u} \, dx && \text{Substitute: } u = \cos x. \\
&= -\ln|u| + C && \text{Apply Log Rule.} \\
&= -\ln|\cos x| + C. && \text{Back-substitute.}
\end{aligned}$$

Try It Exploration A Exploration B

Example 8 uses a trigonometric identity to derive an integration rule for the tangent function. The next example takes a rather unusual step (multiplying and dividing by the same quantity) to derive an integration rule for the secant function.

EXAMPLE 9 Derivation of the Secant Formula

Find $\int \sec x \, dx$.

Solution Consider the following procedure.

$$\begin{aligned}
\int \sec x \, dx &= \int \sec x \left(\frac{\sec x + \tan x}{\sec x + \tan x} \right) dx \\
&= \int \frac{\sec^2 x + \sec x \tan x}{\sec x + \tan x} \, dx
\end{aligned}$$

Letting u be the denominator of this quotient produces

$$u = \sec x + \tan x \quad \Rightarrow \quad u' = \sec x \tan x + \sec^2 x.$$

So, you can conclude that

$$\begin{aligned}
\int \sec x \, dx &= \int \frac{\sec^2 x + \sec x \tan x}{\sec x + \tan x} \, dx && \text{Rewrite integrand.} \\
&= \int \frac{u'}{u} \, dx && \text{Substitute: } u = \sec x + \tan x. \\
&= \ln|u| + C && \text{Apply Log Rule.} \\
&= \ln|\sec x + \tan x| + C. && \text{Back-substitute.}
\end{aligned}$$

Try It Exploration A Video

With the results of Examples 8 and 9, you now have integration formulas for $\sin x$, $\cos x$, $\tan x$, and $\sec x$. All six trigonometric rules are summarized below.

NOTE Using trigonometric identities and properties of logarithms, you could rewrite these six integration rules in other forms. For instance, you could write

$$\int \csc u\, du = \ln|\csc u - \cot u| + C.$$

(See Exercises 83–86.)

Integrals of the Six Basic Trigonometric Functions

$$\int \sin u\, du = -\cos u + C \qquad \int \cos u\, du = \sin u + C$$

$$\int \tan u\, du = -\ln|\cos u| + C \qquad \int \cot u\, du = \ln|\sin u| + C$$

$$\int \sec u\, du = \ln|\sec u + \tan u| + C \qquad \int \csc u\, du = -\ln|\csc u + \cot u| + C$$

EXAMPLE 10 Integrating Trigonometric Functions

Evaluate $\displaystyle\int_0^{\pi/4} \sqrt{1 + \tan^2 x}\, dx$.

Solution Using $1 + \tan^2 x = \sec^2 x$, you can write

$$\begin{aligned}
\int_0^{\pi/4} \sqrt{1 + \tan^2 x}\, dx &= \int_0^{\pi/4} \sqrt{\sec^2 x}\, dx \\
&= \int_0^{\pi/4} \sec x\, dx && \sec x \ge 0 \text{ for } 0 \le x \le \frac{\pi}{4}. \\
&= \ln|\sec x + \tan x|\Big]_0^{\pi/4} \\
&= \ln(\sqrt{2} + 1) - \ln 1 \\
&\approx 0.881.
\end{aligned}$$

Try It Exploration A

The editable graph feature below allows you to edit the graph of a function.

Editable Graph

EXAMPLE 11 Finding an Average Value

Find the average value of $f(x) = \tan x$ on the interval $\left[0, \dfrac{\pi}{4}\right]$.

Solution

$$\begin{aligned}
\text{Average value} &= \frac{1}{(\pi/4) - 0} \int_0^{\pi/4} \tan x\, dx && \text{Average value} = \frac{1}{b-a}\int_a^b f(x)\, dx \\
&= \frac{4}{\pi}\int_0^{\pi/4} \tan x\, dx && \text{Simplify.} \\
&= \frac{4}{\pi}\Big[-\ln|\cos x|\Big]_0^{\pi/4} && \text{Integrate.} \\
&= -\frac{4}{\pi}\left[\ln\left(\frac{\sqrt{2}}{2}\right) - \ln(1)\right] \\
&= -\frac{4}{\pi}\ln\left(\frac{\sqrt{2}}{2}\right) \\
&\approx 0.441
\end{aligned}$$

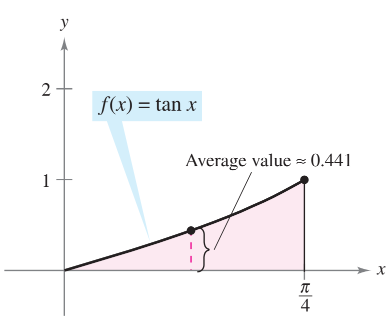


Figure 5.9

The average value is about 0.441, as shown in Figure 5.9.

Try It Exploration A

Section 5.3 Inverse Functions

- **Verify that one function is the inverse function of another function.**
- **Determine whether a function has an inverse function.**
- **Find the derivative of an inverse function.**

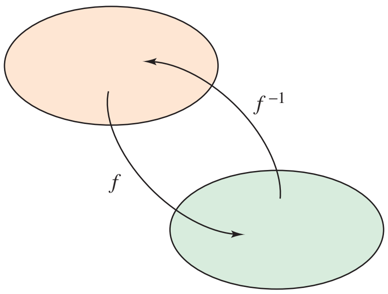
Inverse Functions

Recall from Section P.3 that a function can be represented by a set of ordered pairs. For instance, the function $f(x) = x + 3$ from $A = \{1, 2, 3, 4\}$ to $B = \{4, 5, 6, 7\}$ can be written as

$$f: \{(1, 4), (2, 5), (3, 6), (4, 7)\}.$$

By interchanging the first and second coordinates of each ordered pair, you can form the **inverse function** of f. This function is denoted by f^{-1}. It is a function from B to A, and can be written as

$$f^{-1}: \{(4, 1), (5, 2), (6, 3), (7, 4)\}.$$



Domain of $f =$ range of f^{-1}
Domain of $f^{-1} =$ range of f
Figure 5.10

Note that the domain of f is equal to the range of f^{-1}, and vice versa, as shown in Figure 5.10. The functions f and f^{-1} have the effect of "undoing" each other. That is, when you form the composition of f with f^{-1} or the composition of f^{-1} with f, you obtain the identity function.

$$f(f^{-1}(x)) = x \qquad \text{and} \qquad f^{-1}(f(x)) = x$$

EXPLORATION

Finding Inverse Functions Explain how to "undo" each of the following functions. Then use your explanation to write the inverse function of f.

a. $f(x) = x - 5$

b. $f(x) = 6x$

c. $f(x) = \dfrac{x}{2}$

d. $f(x) = 3x + 2$

e. $f(x) = x^3$

f. $f(x) = 4(x - 2)$

Use a graphing utility to graph each function and its inverse function in the same "square" viewing window. What observation can you make about each pair of graphs?

Definition of Inverse Function

A function g is the **inverse function** of the function f if

$$f(g(x)) = x \quad \text{for each } x \text{ in the domain of } g$$

and

$$g(f(x)) = x \quad \text{for each } x \text{ in the domain of } f.$$

The function g is denoted by f^{-1} (read "f inverse").

NOTE Although the notation used to denote an inverse function resembles *exponential notation*, it is a different use of -1 as a superscript. That is, in general, $f^{-1}(x) \neq 1/f(x)$.

Here are some important observations about inverse functions.

1. If g is the inverse function of f, then f is the inverse function of g.
2. The domain of f^{-1} is equal to the range of f, and the range of f^{-1} is equal to the domain of f.
3. A function need not have an inverse function, but if it does, the inverse function is unique (see Exercise 99).

You can think of f^{-1} as undoing what has been done by f. For example, subtraction can be used to undo addition, and division can be used to undo multiplication. Use the definition of an inverse function to check the following.

$$f(x) = x + c \qquad \text{and} \qquad f^{-1}(x) = x - c \qquad \text{are inverse functions of each other.}$$

$$f(x) = cx \qquad \text{and} \qquad f^{-1}(x) = \frac{x}{c},\ c \neq 0, \qquad \text{are inverse functions of each other.}$$

EXAMPLE 1 Verifying Inverse Functions

Show that the functions are inverse functions of each other.

$$f(x) = 2x^3 - 1 \qquad \text{and} \qquad g(x) = \sqrt[3]{\frac{x+1}{2}}$$

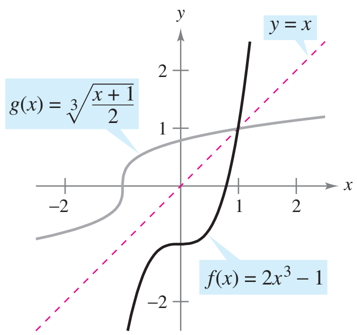
Solution Because the domains and ranges of both f and g consist of all real numbers, you can conclude that both composite functions exist for all x. The composition of f with g is given by

$$\begin{aligned} f(g(x)) &= 2\left(\sqrt[3]{\frac{x+1}{2}}\right)^3 - 1 \\ &= 2\left(\frac{x+1}{2}\right) - 1 \\ &= x + 1 - 1 \\ &= x. \end{aligned}$$

The composition of g with f is given by

$$\begin{aligned} g(f(x)) &= \sqrt[3]{\frac{(2x^3 - 1) + 1}{2}} \\ &= \sqrt[3]{\frac{2x^3}{2}} \\ &= \sqrt[3]{x^3} \\ &= x. \end{aligned}$$

Because $f(g(x)) = x$ and $g(f(x)) = x$, you can conclude that f and g are inverse functions of each other (see Figure 5.11).



f and g are inverse functions of each other.
Figure 5.11

Editable Graph

Try It | Exploration A | Exploration B | Exploration C

STUDY TIP In Example 1, try comparing the functions f and g verbally.

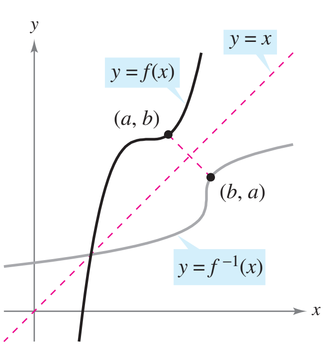
For f: First cube x, then multiply by 2, then subtract 1.

For g: First add 1, then divide by 2, then take the cube root.

Do you see the "undoing pattern"?

In Figure 5.11, the graphs of f and $g = f^{-1}$ appear to be mirror images of each other with respect to the line $y = x$. The graph of f^{-1} is a **reflection** of the graph of f in the line $y = x$. This idea is generalized in the following theorem.

Animation



The graph of f^{-1} is a reflection of the graph of f in the line $y = x$.
Figure 5.12

THEOREM 5.6 Reflective Property of Inverse Functions

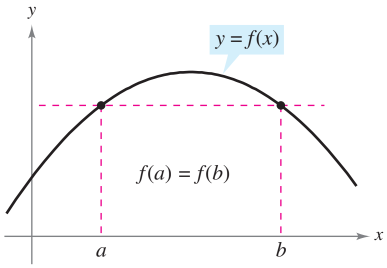
The graph of f contains the point (a, b) if and only if the graph of f^{-1} contains the point (b, a).

Proof If (a, b) is on the graph of f, then $f(a) = b$ and you can write

$$f^{-1}(b) = f^{-1}(f(a)) = a.$$

So, (b, a) is on the graph of f^{-1}, as shown in Figure 5.12. A similar argument will prove the theorem in the other direction.

Technology | Technology



If a horizontal line intersects the graph of f twice, then f is not one-to-one.
Figure 5.13

Existence of an Inverse Function

Not every function has an inverse function, and Theorem 5.6 suggests a graphical test for those that do—the **Horizontal Line Test** for an inverse function. This test states that a function f has an inverse function if and only if every horizontal line intersects the graph of f at most once (see Figure 5.13). The following theorem formally states why the horizontal line test is valid. (Recall from Section 3.3 that a function is *strictly monotonic* if it is either increasing on its entire domain or decreasing on its entire domain.)

THEOREM 5.7 The Existence of an Inverse Function

1. A function has an inverse function if and only if it is one-to-one.
2. If f is strictly monotonic on its entire domain, then it is one-to-one and therefore has an inverse function.

Proof To prove the second part of the theorem, recall from Section P.3 that f is one-to-one if for x_1 and x_2 in its domain

$$f(x_1) = f(x_2) \quad \Rightarrow \quad x_1 = x_2.$$

The *contrapositive* of this implication is logically equivalent and states that

$$x_1 \neq x_2 \quad \Rightarrow \quad f(x_1) \neq f(x_2).$$

Now, choose x_1 and x_2 in the domain of f. If $x_1 \neq x_2$, then, because f is strictly monotonic, it follows that either

$$f(x_1) < f(x_2) \qquad \text{or} \qquad f(x_1) > f(x_2).$$

In either case, $f(x_1) \neq f(x_2)$. So, f is one-to-one on the interval. The proof of the first part of the theorem is left as an exercise (see Exercise 100).

EXAMPLE 2 The Existence of an Inverse Function

Which of the functions has an inverse function?

a. $f(x) = x^3 + x - 1$ **b.** $f(x) = x^3 - x + 1$

Solution

a. From the graph of f shown in Figure 5.14(a), it appears that f is increasing over its entire domain. To verify this, note that the derivative, $f'(x) = 3x^2 + 1$, is positive for all real values of x. So, f is strictly monotonic and it must have an inverse function.

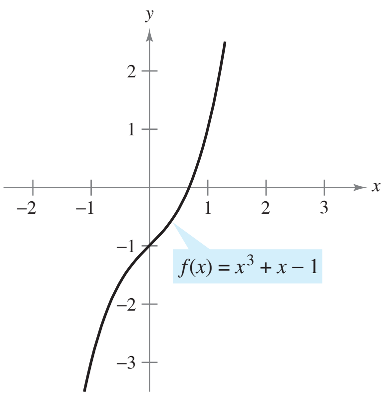
b. From the graph of f shown in Figure 5.14(b), you can see that the function does not pass the horizontal line test. In other words, it is not one-to-one. For instance, f has the same value when $x = -1$, 0, and 1.

$$f(-1) = f(1) = f(0) = 1 \qquad \text{Not one-to-one}$$

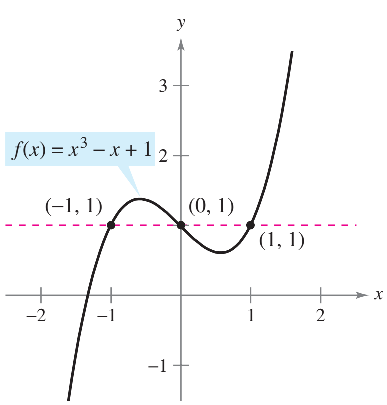
So, by Theorem 5.7, f does not have an inverse function.

Try It **Exploration A**

NOTE Often it is easier to prove that a function *has* an inverse function than to find the inverse function. For instance, it would be difficult algebraically to find the inverse function of the function in Example 2(a).



(a) Because f is increasing over its entire domain, it has an inverse function.



(b) Because f is not one-to-one, it does not have an inverse function.
Figure 5.14

The following guidelines suggest a procedure for finding an inverse function.

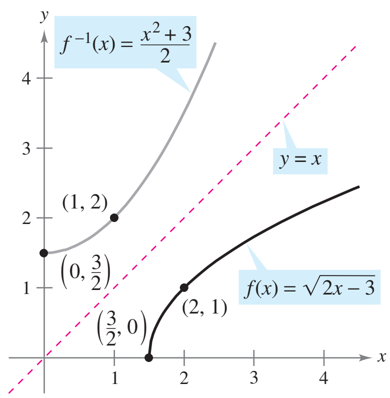
Guidelines for Finding an Inverse Function

1. Use Theorem 5.7 to determine whether the function given by $y = f(x)$ has an inverse function.
2. Solve for x as a function of y: $x = g(y) = f^{-1}(y)$.
3. Interchange x and y. The resulting equation is $y = f^{-1}(x)$.
4. Define the domain of f^{-1} to be the range of f.
5. Verify that $f(f^{-1}(x)) = x$ and $f^{-1}(f(x)) = x$.

EXAMPLE 3 Finding an Inverse Function

Find the inverse function of

$$f(x) = \sqrt{2x - 3}.$$



The domain of f^{-1}, $[0, \infty)$, is the range of f.
Figure 5.15

Editable Graph

Solution The function has an inverse function because it is increasing on its entire domain (see Figure 5.15). To find an equation for the inverse function, let $y = f(x)$ and solve for x in terms of y.

$$\sqrt{2x - 3} = y \qquad \text{Let } y = f(x).$$

$$2x - 3 = y^2 \qquad \text{Square each side.}$$

$$x = \frac{y^2 + 3}{2} \qquad \text{Solve for } x.$$

$$y = \frac{x^2 + 3}{2} \qquad \text{Interchange } x \text{ and } y.$$

$$f^{-1}(x) = \frac{x^2 + 3}{2} \qquad \text{Replace } y \text{ by } f^{-1}(x).$$

The domain of f^{-1} is the range of f, which is $[0, \infty)$. You can verify this result as shown.

$$f(f^{-1}(x)) = \sqrt{2\left(\frac{x^2 + 3}{2}\right) - 3} = \sqrt{x^2} = x, \quad x \geq 0$$

$$f^{-1}(f(x)) = \frac{\left(\sqrt{2x - 3}\right)^2 + 3}{2} = \frac{2x - 3 + 3}{2} = x, \quad x \geq \frac{3}{2}$$

Try It

Exploration A

NOTE Remember that any letter can be used to represent the independent variable. So,

$$f^{-1}(y) = \frac{y^2 + 3}{2}$$

$$f^{-1}(x) = \frac{x^2 + 3}{2}$$

$$f^{-1}(s) = \frac{s^2 + 3}{2}$$

all represent the same function.

Theorem 5.7 is useful in the following type of problem. Suppose you are given a function that is *not* one-to-one on its domain. By restricting the domain to an interval on which the function is strictly monotonic, you can conclude that the new function *is* one-to-one on the restricted domain.

EXAMPLE 4 Testing Whether a Function Is One-to-One

Show that the sine function

$$f(x) = \sin x$$

is not one-to-one on the entire real line. Then show that $[-\pi/2, \pi/2]$ is the largest interval, centered at the origin, for which f is strictly monotonic.

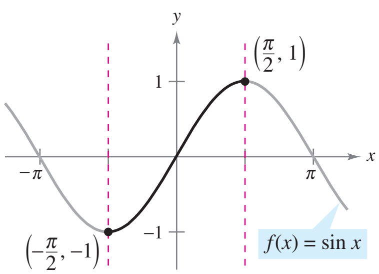
Solution It is clear that f is not one-to-one, because many different x-values yield the same y-value. For instance,

$$\sin(0) = 0 = \sin(\pi).$$

Moreover, f is increasing on the open interval $(-\pi/2, \pi/2)$, because its derivative

$$f'(x) = \cos x$$

is positive there. Finally, because the left and right endpoints correspond to relative extrema of the sine function, you can conclude that f is increasing on the closed interval $[-\pi/2, \pi/2]$ *and* that in any larger interval the function is not strictly monotonic (see Figure 5.16).



f is one-to-one on the interval $[-\pi/2, \pi/2]$.
Figure 5.16

Editable Graph

Try It | Exploration A | Open Exploration

Derivative of an Inverse Function

The next two theorems discuss the derivative of an inverse function. The reasonableness of Theorem 5.8 follows from the reflective property of inverse functions as shown in Figure 5.12. Proofs of the two theorems are given in Appendix A.

THEOREM 5.8 Continuity and Differentiability of Inverse Functions

Let f be a function whose domain is an interval I. If f has an inverse function, then the following statements are true.

1. If f is continuous on its domain, then f^{-1} is continuous on its domain.
2. If f is increasing on its domain, then f^{-1} is increasing on its domain.
3. If f is decreasing on its domain, then f^{-1} is decreasing on its domain.
4. If f is differentiable at c and $f'(c) \neq 0$, then f^{-1} is differentiable at $f(c)$.

THEOREM 5.9 The Derivative of an Inverse Function

Let f be a function that is differentiable on an interval I. If f has an inverse function g, then g is differentiable at any x for which $f'(g(x)) \neq 0$. Moreover,

$$g'(x) = \frac{1}{f'(g(x))}, \quad f'(g(x)) \neq 0.$$

EXPLORATION

Graph the inverse functions

$$f(x) = x^3$$

and

$$g(x) = x^{1/3}.$$

Calculate the slope of f at $(1, 1)$, $(2, 8)$, and $(3, 27)$, and the slope of g at $(1, 1)$, $(8, 2)$, and $(27, 3)$. What do you observe? What happens at $(0, 0)$?

EXAMPLE 5 Evaluating the Derivative of an Inverse Function

Let $f(x) = \frac{1}{4}x^3 + x - 1$.

a. What is the value of $f^{-1}(x)$ when $x = 3$?

b. What is the value of $(f^{-1})'(x)$ when $x = 3$?

Solution Notice that f is one-to-one and therefore has an inverse function.

a. Because $f(x) = 3$ when $x = 2$, you know that $f^{-1}(3) = 2$.

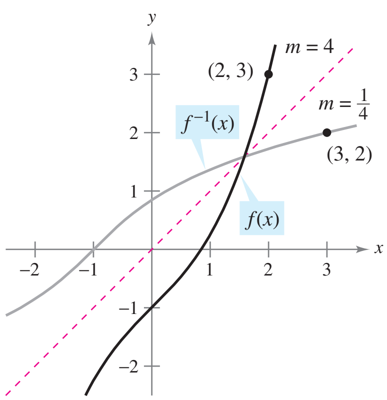
b. Because the function f is differentiable and has an inverse function, you can apply Theorem 5.9 (with $g = f^{-1}$) to write

$$(f^{-1})'(3) = \frac{1}{f'(f^{-1}(3))} = \frac{1}{f'(2)}.$$

Moreover, using $f'(x) = \frac{3}{4}x^2 + 1$, you can conclude that

$$(f^{-1})'(3) = \frac{1}{f'(2)} = \frac{1}{\frac{3}{4}(2^2) + 1} = \frac{1}{4}.$$

Try It | Exploration A | Exploration B | Video



The graphs of the inverse functions f and f^{-1} have reciprocal slopes at points (a, b) and (b, a).
Figure 5.17

Editable Graph

In Example 5, note that at the point $(2, 3)$ the slope of the graph of f is 4 and at the point $(3, 2)$ the slope of the graph of f^{-1} is $\frac{1}{4}$ (see Figure 5.17). This reciprocal relationship (which follows from Theorem 5.9) can be written as shown below. If $y = g(x) = f^{-1}(x)$, then $f(y) = x$ and $f'(y) = \frac{dx}{dy}$. Theorem 5.9 says that

$$g'(x) = \frac{dy}{dx} = \frac{1}{f'(g(x))} = \frac{1}{f'(y)} = \frac{1}{(dx/dy)}.$$

So, $\boxed{\frac{dy}{dx} = \frac{1}{dx/dy}.}$

EXAMPLE 6 Graphs of Inverse Functions Have Reciprocal Slopes

Let $f(x) = x^2$ (for $x \geq 0$) and let $f^{-1}(x) = \sqrt{x}$. Show that the slopes of the graphs of f and f^{-1} are reciprocals at each of the following points.

a. $(2, 4)$ and $(4, 2)$ **b.** $(3, 9)$ and $(9, 3)$

Solution The derivatives of f and f^{-1} are given by

$$f'(x) = 2x \quad \text{and} \quad (f^{-1})'(x) = \frac{1}{2\sqrt{x}}.$$

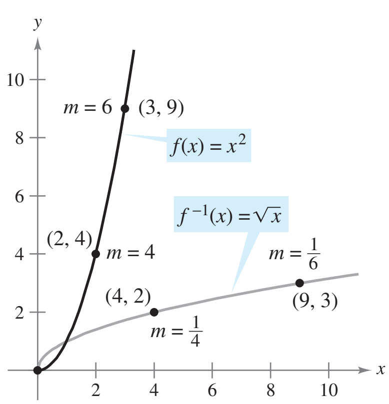
a. At $(2, 4)$, the slope of the graph of f is $f'(2) = 2(2) = 4$. At $(4, 2)$, the slope of the graph of f^{-1} is

$$(f^{-1})'(4) = \frac{1}{2\sqrt{4}} = \frac{1}{2(2)} = \frac{1}{4}.$$

b. At $(3, 9)$, the slope of the graph of f is $f'(3) = 2(3) = 6$. At $(9, 3)$, the slope of the graph of f^{-1} is

$$(f^{-1})'(9) = \frac{1}{2\sqrt{9}} = \frac{1}{2(3)} = \frac{1}{6}.$$

So, in both cases, the slopes are reciprocals, as shown in Figure 5.18.



At $(0, 0)$, the derivative of f is 0, and the derivative of f^{-1} does not exist.
Figure 5.18

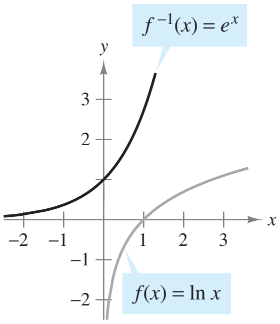
Editable Graph

Try It | Exploration A | Exploration B

Section 5.4 Exponential Functions: Differentiation and Integration

- **Develop properties of the natural exponential function.**
- **Differentiate natural exponential functions.**
- **Integrate natural exponential functions.**

The Natural Exponential Function



The inverse function of the natural logarithmic function is the natural exponential function.
Figure 5.19

The function $f(x) = \ln x$ is increasing on its entire domain, and therefore it has an inverse function f^{-1}. The domain of f^{-1} is the set of all reals, and the range is the set of positive reals, as shown in Figure 5.19. So, for any real number x,

$$f(f^{-1}(x)) = \ln[f^{-1}(x)] = x. \qquad x \text{ is any real number.}$$

If x happens to be rational, then

$$\ln(e^x) = x \ln e = x(1) = x. \qquad x \text{ is a rational number.}$$

Because the natural logarithmic function is one-to-one, you can conclude that $f^{-1}(x)$ and e^x agree for *rational* values of x. The following definition extends the meaning of e^x to include *all* real values of x.

Definition of the Natural Exponential Function

The inverse function of the natural logarithmic function $f(x) = \ln x$ is called the **natural exponential function** and is denoted by

$$f^{-1}(x) = e^x.$$

That is,

$$y = e^x \quad \text{if and only if} \quad x = \ln y.$$

The Number *e*

The symbol e was first used by mathematician Leonhard Euler to represent the base of natural logarithms in a letter to another mathematician, Christian Goldbach, in 1731.

The inverse relationship between the natural logarithmic function and the natural exponential function can be summarized as follows.

$$\ln(e^x) = x \quad \text{and} \quad e^{\ln x} = x \qquad \text{Inverse relationship}$$

EXAMPLE 1 Solving Exponential Equations

Solve $7 = e^{x+1}$.

Solution You can convert from exponential form to logarithmic form by *taking the natural logarithim of each side* of the equation.

$$
\begin{aligned}
7 &= e^{x+1} && \text{Write original equation.}\\
\ln 7 &= \ln(e^{x+1}) && \text{Take natural logarithm of each side.}\\
\ln 7 &= x + 1 && \text{Apply inverse property.}\\
-1 + \ln 7 &= x && \text{Solve for } x.\\
0.946 &\approx x && \text{Use a calculator.}
\end{aligned}
$$

Check this solution in the original equation.

Try It | Exploration A | Exploration B

EXAMPLE 2 Solving a Logarithmic Equation

Solve $\ln(2x - 3) = 5$.

Solution To convert from logarithmic form to exponential form, you can *exponentiate each side* of the logarithmic equation.

$$\ln(2x - 3) = 5 \qquad \text{Write original equation.}$$
$$e^{\ln(2x-3)} = e^5 \qquad \text{Exponentiate each side.}$$
$$2x - 3 = e^5 \qquad \text{Apply inverse property.}$$
$$x = \tfrac{1}{2}(e^5 + 3) \qquad \text{Solve for } x.$$
$$x \approx 75.707 \qquad \text{Use a calculator.}$$

Try It   Exploration A

The editable graph feature below allows you to edit the graph of a function.

Editable Graph

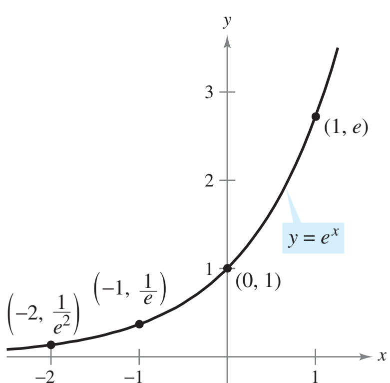
The familiar rules for operating with rational exponents can be extended to the natural exponential function, as shown in the following theorem.

THEOREM 5.10 Operations with Exponential Functions

Let a and b be any real numbers.

1. $e^a e^b = e^{a+b}$

2. $\dfrac{e^a}{e^b} = e^{a-b}$



The natural exponential function is increasing, and its graph is concave upward.
Figure 5.20

Proof To prove Property 1, you can write

$$\begin{aligned}\ln(e^a e^b) &= \ln(e^a) + \ln(e^b)\\ &= a + b\\ &= \ln(e^{a+b}).\end{aligned}$$

Because the natural logarithmic function is one-to-one, you can conclude that

$$e^a e^b = e^{a+b}.$$

The proof of the second property is left to you (see Exercise 129).

In Section 5.3, you learned that an inverse function f^{-1} shares many properties with f. So, the natural exponential function inherits the following properties from the natural logarithmic function (see Figure 5.20).

Properties of the Natural Exponential Function

1. The domain of $f(x) = e^x$ is $(-\infty, \infty)$, and the range is $(0, \infty)$.

2. The function $f(x) = e^x$ is continuous, increasing, and one-to-one on its entire domain.

3. The graph of $f(x) = e^x$ is concave upward on its entire domain.

4. $\displaystyle\lim_{x\to-\infty} e^x = 0$ and $\displaystyle\lim_{x\to\infty} e^x = \infty$

Derivatives of Exponential Functions

One of the most intriguing (and useful) characteristics of the natural exponential function is that *it is its own derivative*. In other words, it is a solution to the differential equation $y' = y$. This result is stated in the next theorem.

FOR FURTHER INFORMATION To find out about derivatives of exponential functions of order 1/2, see the article "A Child's Garden of Fractional Derivatives" by Marcia Kleinz and Thomas J. Osler in *The College Mathematics Journal*.

MathArticle

THEOREM 5.11 Derivative of the Natural Exponential Function

Let u be a differentiable function of x.

1. $\dfrac{d}{dx}[e^x] = e^x$

2. $\dfrac{d}{dx}[e^u] = e^u \dfrac{du}{dx}$

Proof To prove Property 1, use the fact that $\ln e^x = x$, and differentiate each side of the equation.

$$\ln e^x = x \qquad \text{Definition of exponential function}$$

$$\frac{d}{dx}[\ln e^x] = \frac{d}{dx}[x] \qquad \text{Differentiate each side with respect to } x.$$

$$\frac{1}{e^x}\frac{d}{dx}[e^x] = 1$$

$$\frac{d}{dx}[e^x] = e^x$$

The derivative of e^u follows from the Chain Rule.

NOTE You can interpret this theorem geometrically by saying that the slope of the graph of $f(x) = e^x$ at any point (x, e^x) is equal to the y-coordinate of the point.

EXAMPLE 3 Differentiating Exponential Functions

a. $\dfrac{d}{dx}[e^{2x-1}] = e^u \dfrac{du}{dx} = 2e^{2x-1}$ $\qquad u = 2x - 1$

b. $\dfrac{d}{dx}[e^{-3/x}] = e^u \dfrac{du}{dx} = \left(\dfrac{3}{x^2}\right)e^{-3/x} = \dfrac{3e^{-3/x}}{x^2}$ $\qquad u = -\dfrac{3}{x}$

Try It Exploration A Video

EXAMPLE 4 Locating Relative Extrema

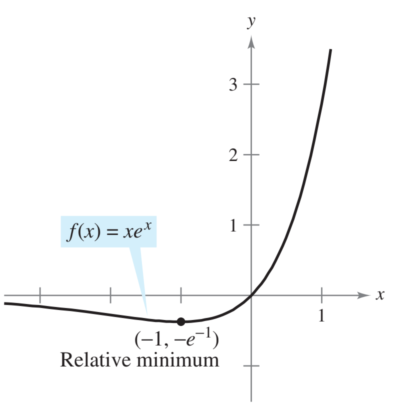
Find the relative extrema of $f(x) = xe^x$.

Solution The derivative of f is given by

$$f'(x) = x(e^x) + e^x(1) \qquad \text{Product Rule}$$
$$= e^x(x + 1).$$

Because e^x is never 0, the derivative is 0 only when $x = -1$. Moreover, by the First Derivative Test, you can determine that this corresponds to a relative minimum, as shown in Figure 5.21. Because the derivative $f'(x) = e^x(x + 1)$ is defined for all x, there are no other critical points.

Try It Exploration A



The derivative of f changes from negative to positive at $x = -1$.
Figure 5.21

Editable Graph

EXAMPLE 5 The Standard Normal Probability Density Function

Show that the *standard normal probability density function*

$$f(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}$$

has points of inflection when $x = \pm 1$.

Solution To locate possible points of inflection, find the x-values for which the second derivative is 0.

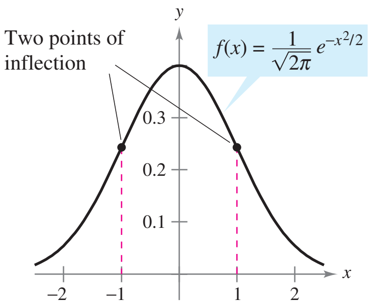
$$f(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2} \quad \text{Write original function.}$$

$$f'(x) = \frac{1}{\sqrt{2\pi}}(-x)e^{-x^2/2} \quad \text{First derivative}$$

$$f''(x) = \frac{1}{\sqrt{2\pi}}[(-x)(-x)e^{-x^2/2} + (-1)e^{-x^2/2}] \quad \text{Product Rule}$$

$$= \frac{1}{\sqrt{2\pi}}(e^{-x^2/2})(x^2 - 1) \quad \text{Second derivative}$$

So, $f''(x) = 0$ when $x = \pm 1$, and you can apply the techniques of Chapter 3 to conclude that these values yield the two points of inflection shown in Figure 5.22.



The bell-shaped curve given by a standard normal probability density function
Figure 5.22

Editable Graph

Try It Exploration A Open Exploration

NOTE The general form of a normal probability density function (whose mean is 0) is given by

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}e^{-x^2/2\sigma^2}$$

where σ is the standard deviation (σ is the lowercase Greek letter sigma). This “bell-shaped curve” has points of inflection when $x = \pm\sigma$.

EXAMPLE 6 Shares Traded

The number y of shares traded (in millions) on the New York Stock Exchange from 1990 through 2002 can be modeled by

$$y = 36{,}663e^{0.1902t}$$

where t represents the year, with $t = 0$ corresponding to 1990. At what rate was the number of shares traded changing in 1998? *(Source: New York Stock Exchange, Inc.)*

Solution The derivative of the given model is

$$y' = (0.1902)(36{,}663)e^{0.1902t}$$
$$\approx 6973e^{0.1902t}.$$

By evaluating the derivative when $t = 8$, you can conclude that the rate of change in 1998 was about

31,933 million shares per year.

The graph of this model is shown in Figure 5.23.

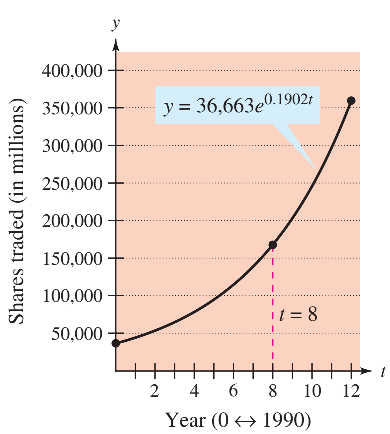


Figure 5.23

Editable Graph

Try It Exploration A

Integrals of Exponential Functions

Each differentiation formula in Theorem 5.11 has a corresponding integration formula.

THEOREM 5.12 Integration Rules for Exponential Functions

Let u be a differentiable function of x.

1. $\displaystyle\int e^x\,dx = e^x + C$ **2.** $\displaystyle\int e^u\,du = e^u + C$

EXAMPLE 7 Integrating Exponential Functions

Find $\displaystyle\int e^{3x+1}\,dx$.

Solution If you let $u = 3x + 1$, then $du = 3\,dx$.

$$\begin{aligned}
\int e^{3x+1}dx &= \frac{1}{3}\int e^{3x+1}(3)\,dx && \text{Multiply and divide by 3.}\\
&= \frac{1}{3}\int e^{u}\,du && \text{Substitute: } u = 3x+1.\\
&= \frac{1}{3}e^u + C && \text{Apply Exponential Rule.}\\
&= \frac{e^{3x+1}}{3} + C && \text{Back-substitute.}
\end{aligned}$$

Try It | Exploration A | Video

NOTE In Example 7, the missing *constant* factor 3 was introduced to create $du = 3\,dx$. However, remember that you cannot introduce a missing *variable* factor in the integrand. For instance,

$$\int e^{-x^2}\,dx \neq \frac{1}{x}\int e^{-x^2}\,(x\,dx).$$

EXAMPLE 8 Integrating Exponential Functions

Find $\displaystyle\int 5xe^{-x^2}\,dx$.

Solution If you let $u = -x^2$, then $du = -2x\,dx$ or $x\,dx = -du/2$.

$$\begin{aligned}
\int 5xe^{-x^2}dx &= \int 5e^{-x^2}(x\,dx) && \text{Regroup integrand.}\\
&= \int 5e^{u}\left(-\frac{du}{2}\right) && \text{Substitute: } u = -x^2.\\
&= -\frac{5}{2}\int e^u\,du && \text{Constant Multiple Rule}\\
&= -\frac{5}{2}e^u + C && \text{Apply Exponential Rule.}\\
&= -\frac{5}{2}e^{-x^2} + C && \text{Back-substitute.}
\end{aligned}$$

Try It | Exploration A

EXAMPLE 9 Integrating Exponential Functions

a. $$\int \frac{e^{1/x}}{x^2}\,dx = -\int \underbrace{e^{1/x}}_{e^u}\underbrace{\left(-\frac{1}{x^2}\right)dx}_{du} \qquad u = \frac{1}{x}$$
$$= -e^{1/x} + C$$

b. $$\int \sin x\, e^{\cos x}\,dx = -\int \underbrace{e^{\cos x}}_{e^u}\underbrace{(-\sin x\,dx)}_{du} \qquad u = \cos x$$
$$= -e^{\cos x} + C$$

Try It Exploration A

EXAMPLE 10 Finding Areas Bounded by Exponential Functions

Evaluate each definite integral.

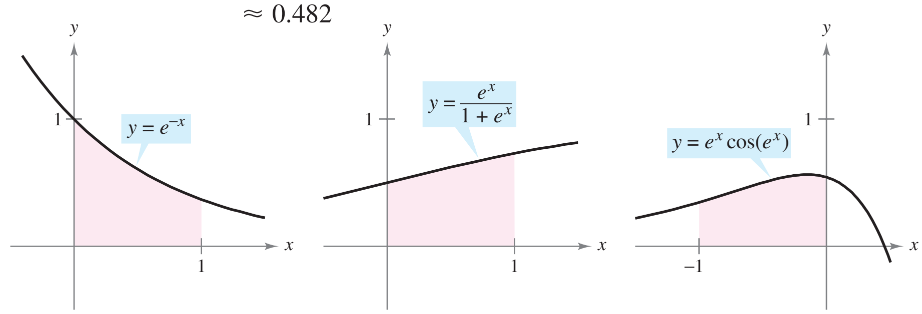
a. $\displaystyle\int_0^1 e^{-x}\,dx$ **b.** $\displaystyle\int_0^1 \frac{e^x}{1+e^x}\,dx$ **c.** $\displaystyle\int_{-1}^0 [e^x\cos(e^x)]\,dx$

Solution

a. $$\int_0^1 e^{-x}\,dx = -e^{-x}\Big]_0^1$$ See Figure 5.24(a).
$$= -e^{-1} - (-1)$$
$$= 1 - \frac{1}{e}$$
$$\approx 0.632$$

b. $$\int_0^1 \frac{e^x}{1+e^x}\,dx = \ln(1+e^x)\Big]_0^1$$ See Figure 5.24(b).
$$= \ln(1+e) - \ln 2$$
$$\approx 0.620$$

c. $$\int_{-1}^0 [e^x\cos(e^x)]\,dx = \sin(e^x)\Big]_{-1}^0$$ See Figure 5.24(c).
$$= \sin 1 - \sin(e^{-1})$$
$$\approx 0.482$$



(a) (b) (c)

Editable Graph Editable Graph Editable Graph

Figure 5.24

Try It Exploration A

Section 5.5 Bases Other Than e and Applications

- **Define exponential functions that have bases other than *e*.**
- **Differentiate and integrate exponential functions that have bases other than *e*.**
- **Use exponential functions to model compound interest and exponential growth.**

Bases Other than e

The **base** of the natural exponential function is e. This "natural" base can be used to assign a meaning to a general base a.

> **Definition of Exponential Function to Base *a***
>
> If a is a positive real number $(a \neq 1)$ and x is any real number, then the **exponential function to the base *a*** is denoted by a^x and is defined by
>
> $$a^x = e^{(\ln a)x}.$$
>
> If $a = 1$, then $y = 1^x = 1$ is a constant function.

These functions obey the usual laws of exponents. For instance, here are some familiar properties.

1. $a^0 = 1$ **2.** $a^x a^y = a^{x+y}$

3. $\dfrac{a^x}{a^y} = a^{x-y}$ **4.** $(a^x)^y = a^{xy}$

When modeling the half-life of a radioactive sample, it is convenient to use $\frac{1}{2}$ as the base of the exponential model.

EXAMPLE 1 Radioactive Half-Life Model

The half-life of carbon-14 is about 5715 years. A sample contains 1 gram of carbon-14. How much will be present in 10,000 years?

Solution Let $t = 0$ represent the present time and let y represent the amount (in grams) of carbon-14 in the sample. Using a base of $\frac{1}{2}$, you can model y by the equation

$$y = \left(\frac{1}{2}\right)^{t/5715}.$$

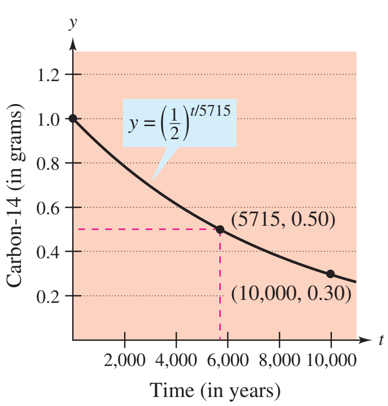
Notice that when $t = 5715$, the amount is reduced to half of the original amount.

$$y = \left(\frac{1}{2}\right)^{5715/5715} = \frac{1}{2}\text{ gram}$$

When $t = 11{,}430$, the amount is reduced to a quarter of the original amount, and so on. To find the amount of carbon-14 after 10,000 years, substitute 10,000 for t.

$$y = \left(\frac{1}{2}\right)^{10{,}000/5715} \approx 0.30\text{ gram}$$

The graph of y is shown in Figure 5.25.



The half-life of carbon-14 is about 5715 years.
Figure 5.25

Editable Graph

Try It

Exploration A

Logarithmic functions to bases other than e can be defined in much the same way as exponential functions to other bases are defined.

NOTE In precalculus, you learned that $\log_a x$ is the value to which a must be raised to produce x. This agrees with the definition given here because

$$\begin{aligned} a^{\log_a x} &= a^{(1/\ln a)\ln x} \\ &= (e^{\ln a})^{(1/\ln a)\ln x} \\ &= e^{(\ln a/\ln a)\ln x} \\ &= e^{\ln x} \\ &= x. \end{aligned}$$

Definition of Logarithmic Function to Base a

If a is a positive real number $(a \neq 1)$ and x is any positive real number, then the **logarithmic function to the base a** is denoted by $\log_a x$ and is defined as

$$\log_a x = \frac{1}{\ln a} \ln x.$$

Logarithmic functions to the base a have properties similar to those of the natural logarithmic function given in Theorem 5.2.

1. $\log_a 1 = 0$ — Log of 1

2. $\log_a xy = \log_a x + \log_a y$ — Log of a product

3. $\log_a x^n = n \log_a x$ — Log of a power

4. $\log_a \dfrac{x}{y} = \log_a x - \log_a y$ — Log of a quotient

From the definitions of the exponential and logarithmic functions to the base a, it follows that $f(x) = a^x$ and $g(x) = \log_a x$ are inverse functions of each other.

Properties of Inverse Functions

1. $y = a^x$ if and only if $x = \log_a y$

2. $a^{\log_a x} = x, \quad$ for $x > 0$

3. $\log_a a^x = x, \quad$ for all x

The logarithmic function to the base 10 is called the **common logarithmic function.** So, for common logarithms, $y = 10^x$ if and only if $x = \log_{10} y$.

EXAMPLE 2 Bases Other Than e

Solve for x in each equation.

a. $3^x = \dfrac{1}{81}$

b. $\log_2 x = -4$

Solution

a. To solve this equation, you can apply the logarithmic function to the base 3 to each side of the equation.

$$\begin{aligned} 3^x &= \frac{1}{81} \\ \log_3 3^x &= \log_3 \frac{1}{81} \\ x &= \log_3 3^{-4} \\ x &= -4 \end{aligned}$$

b. To solve this equation, you can apply the exponential function to the base 2 to each side of the equation.

$$\begin{aligned} \log_2 x &= -4 \\ 2^{\log_2 x} &= 2^{-4} \\ x &= \frac{1}{2^4} \\ x &= \frac{1}{16} \end{aligned}$$

Try It

Exploration A

Differentiation and Integration

To differentiate exponential and logarithmic functions to other bases, you have three options: (1) use the definitions of a^x and $\log_a x$ and differentiate using the rules for the natural exponential and logarithmic functions, (2) use logarithmic differentiation, or (3) use the following differentiation rules for bases other than e.

THEOREM 5.13 Derivatives for Bases Other Than e

Let a be a positive real number $(a \neq 1)$ and let u be a differentiable function of x.

1. $\dfrac{d}{dx}[a^x] = (\ln a)a^x$

2. $\dfrac{d}{dx}[a^u] = (\ln a)a^u \dfrac{du}{dx}$

3. $\dfrac{d}{dx}[\log_a x] = \dfrac{1}{(\ln a)x}$

4. $\dfrac{d}{dx}[\log_a u] = \dfrac{1}{(\ln a)u}\dfrac{du}{dx}$

Video

Proof By definition, $a^x = e^{(\ln a)x}$. So, you can prove the first rule by letting $u = (\ln a)x$ and differentiating with base e to obtain

$$\frac{d}{dx}[a^x] = \frac{d}{dx}[e^{(\ln a)x}] = e^u \frac{du}{dx} = e^{(\ln a)x}(\ln a) = (\ln a)a^x.$$

To prove the third rule, you can write

$$\frac{d}{dx}[\log_a x] = \frac{d}{dx}\left[\frac{1}{\ln a}\ln x\right] = \frac{1}{\ln a}\left(\frac{1}{x}\right) = \frac{1}{(\ln a)x}.$$

The second and fourth rules are simply the Chain Rule versions of the first and third rules.

NOTE These differentiation rules are similar to those for the natural exponential function and natural logarithmic function. In fact, they differ only by the constant factors $\ln a$ and $1/\ln a$. This points out one reason why, for calculus, e is the most convenient base.

EXAMPLE 3 Differentiating Functions to Other Bases

Find the derivative of each function.

a. $y = 2^x$

b. $y = 2^{3x}$

c. $y = \log_{10} \cos x$

Solution

a. $y' = \dfrac{d}{dx}[2^x] = (\ln 2)2^x$

b. $y' = \dfrac{d}{dx}[2^{3x}] = (\ln 2)2^{3x}(3) = (3 \ln 2)2^{3x}$

Try writing 2^{3x} as 8^x and differentiating to see that you obtain the same result.

c. $y' = \dfrac{d}{dx}[\log_{10} \cos x] = \dfrac{-\sin x}{(\ln 10)\cos x} = -\dfrac{1}{\ln 10}\tan x$

Try It Exploration A Technology Video

Occasionally, an integrand involves an exponential function to a base other than e. When this occurs, there are two options: (1) convert to base e using the formula $a^x = e^{(\ln a)x}$ and then integrate, or (2) integrate directly, using the integration formula

$$\int a^x\,dx = \left(\frac{1}{\ln a}\right)a^x + C$$

(which follows from Theorem 5.13).

EXAMPLE 4 Integrating an Exponential Function to Another Base

Find $\int 2^x\,dx$.

Solution

$$\int 2^x\,dx = \frac{1}{\ln 2}2^x + C$$

Try It Exploration A

When the Power Rule, $D_x[x^n] = nx^{n-1}$, was introduced in Chapter 2, the exponent n was required to be a rational number. Now the rule is extended to cover any real value of n. Try to prove this theorem using logarithmic differentiation.

THEOREM 5.14 The Power Rule for Real Exponents

Let n be any real number and let u be a differentiable function of x.

1. $\dfrac{d}{dx}[x^n] = nx^{n-1}$
2. $\dfrac{d}{dx}[u^n] = nu^{n-1}\dfrac{du}{dx}$

Technology

The next example compares the derivatives of four types of functions. Each function uses a different differentiation formula, depending on whether the base and exponent are constants or variables.

EXAMPLE 5 Comparing Variables and Constants

a. $\dfrac{d}{dx}[e^e] = 0$ Constant Rule

b. $\dfrac{d}{dx}[e^x] = e^x$ Exponential Rule

c. $\dfrac{d}{dx}[x^e] = ex^{e-1}$ Power Rule

d. $y = x^x$ Logarithmic differentiation

$$\ln y = \ln x^x$$
$$\ln y = x\ln x$$
$$\frac{y'}{y} = x\left(\frac{1}{x}\right) + (\ln x)(1) = 1 + \ln x$$
$$y' = y(1 + \ln x) = x^x(1 + \ln x)$$

NOTE Be sure you see that there is no simple differentiation rule for calculating the derivative of $y = x^x$. In general, if $y = u(x)^{v(x)}$, you need to use logarithmic differentiation.

Try It Exploration A

Applications of Exponential Functions

n	A
1	\$1080.00
2	\$1081.60
4	\$1082.43
12	\$1083.00
365	\$1083.28

Suppose P dollars is deposited in an account at an annual interest rate r (in decimal form). If interest accumulates in the account, what is the balance in the account at the end of 1 year? The answer depends on the number of times n the interest is compounded according to the formula

$$A = P\left(1 + \frac{r}{n}\right)^n.$$

For instance, the result for a deposit of \$1000 at 8% interest compounded n times a year is shown in the upper table at the left.

As n increases, the balance A approaches a limit. To develop this limit, use the following theorem. To test the reasonableness of this theorem, try evaluating $[(x + 1)/x]^x$ for several values of x, as shown in the lower table at the left. (A proof of this theorem is given in Appendix A.)

x	$\left(\frac{x+1}{x}\right)^x$
10	2.59374
100	2.70481
1000	2.71692
10,000	2.71815
100,000	2.71827
1,000,000	2.71828

THEOREM 5.15 A Limit Involving e

$$\lim_{x\to\infty}\left(1 + \frac{1}{x}\right)^x = \lim_{x\to\infty}\left(\frac{x+1}{x}\right)^x = e$$

Now, let's take another look at the formula for the balance A in an account in which the interest is compounded n times per year. By taking the limit as n approaches infinity, you obtain

$$\begin{aligned}
A &= \lim_{n\to\infty} P\left(1 + \frac{r}{n}\right)^n && \text{Take limit as } n \to \infty.\\
&= P \lim_{n\to\infty}\left[\left(1 + \frac{1}{n/r}\right)^{n/r}\right]^r && \text{Rewrite.}\\
&= P\left[\lim_{x\to\infty}\left(1 + \frac{1}{x}\right)^x\right]^r && \text{Let } x = n/r. \text{ Then } x \to \infty \text{ as } n \to \infty.\\
&= Pe^r. && \text{Apply Theorem 5.15.}
\end{aligned}$$

This limit produces the balance after 1 year of **continuous compounding.** So, for a deposit of \$1000 at 8% interest compounded continuously, the balance at the end of 1 year would be

$$\begin{aligned}
A &= 1000e^{0.08}\\
&\approx \$1083.29.
\end{aligned}$$

These results are summarized below.

Summary of Compound Interest Formulas

Let P = amount of deposit, t = number of years, A = balance after t years, r = annual interest rate (decimal form), and n = number of compoundings per year.

1. Compounded n times per year: $A = P\left(1 + \frac{r}{n}\right)^{nt}$

2. Compounded continuously: $A = Pe^{rt}$

EXAMPLE 6 Comparing Continuous and Quarterly Compounding

A deposit of \$2500 is made in an account that pays an annual interest rate of 5%. Find the balance in the account at the end of 5 years if the interest is compounded (a) quarterly, (b) monthly, and (c) continuously.

Solution

a. $A = P\left(1 + \frac{r}{n}\right)^{nt} = 2500\left(1 + \frac{0.05}{4}\right)^{4(5)}$ Compounded quarterly

$= 2500(1.0125)^{20}$

$\approx \$3205.09$

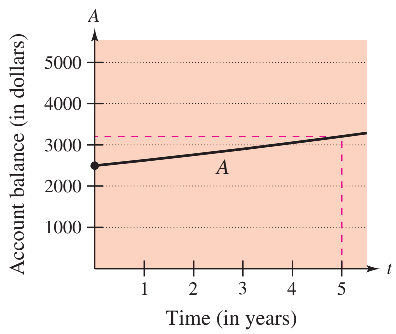
b. $A = P\left(1 + \frac{r}{n}\right)^{nt} = 2500\left(1 + \frac{0.05}{12}\right)^{12(5)}$ Compounded monthly

$\approx 2500(1.0041667)^{60}$

$\approx \$3208.40$

c. $A = Pe^{rt} = 2500\left[e^{0.05(5)}\right]$ Compounded continuously

$= 2500e^{0.25} \approx \$3210.06$



The balance in a savings account grows exponentially.
Figure 5.26

Figure 5.26 shows how the balance increases over the five-year period. Notice that the scale used in the figure does not graphically distinguish among the three types of exponential growth in (a), (b), and (c).

Try It | Exploration A | Open Exploration

EXAMPLE 7 Bacterial Culture Growth

A bacterial culture is growing according to the *logistic growth function*

$$y = \frac{1.25}{1 + 0.25e^{-0.4t}}, \quad t \geq 0$$

where y is the weight of the culture in grams and t is the time in hours. Find the weight of the culture after (a) 0 hours, (b) 1 hour, and (c) 10 hours. (d) What is the limit as t approaches infinity?

Solution

a. When $t = 0$, $y = \frac{1.25}{1 + 0.25e^{-0.4(0)}}$

$= 1$ gram.

b. When $t = 1$, $y = \frac{1.25}{1 + 0.25e^{-0.4(1)}}$

≈ 1.071 grams.

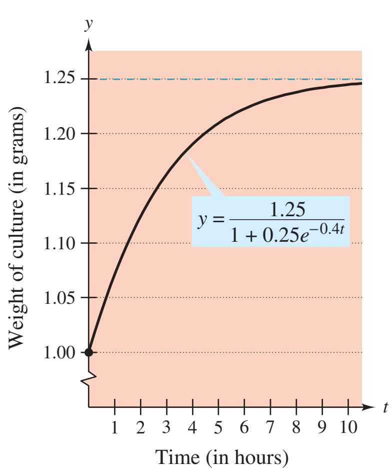
c. When $t = 10$, $y = \frac{1.25}{1 + 0.25e^{-0.4(10)}}$

≈ 1.244 grams.

d. Finally, taking the limit as t approaches infinity, you obtain

$$\lim_{t \to \infty} \frac{1.25}{1 + 0.25e^{-0.4t}} = \frac{1.25}{1 + 0} = 1.25 \text{ grams.}$$

The graph of the function is shown in Figure 5.27.



The limit of the weight of the culture as $t \to \infty$ is 1.25 grams.
Figure 5.27

Editable Graph

Try It | Exploration A

Section 5.6 Inverse Trigonometric Functions: Differentiation

- **Develop properties of the six inverse trigonometric functions.**
- **Differentiate an inverse trigonometric function.**
- **Review the basic differentiation rules for elementary functions.**

Inverse Trigonometric Functions

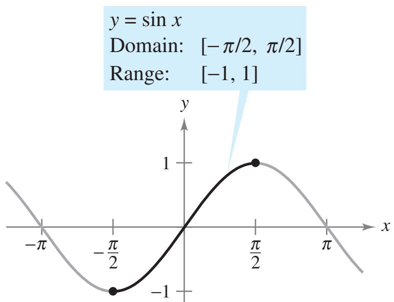
This section begins with a rather surprising statement: *None of the six basic trigonometric functions has an inverse function*. This statement is true because all six trigonometric functions are periodic and therefore are not one-to-one. In this section you will examine these six functions to see whether their domains can be redefined in such a way that they will have inverse functions on the *restricted domains*.

In Example 4 of Section 5.3, you saw that the sine function is increasing (and therefore is one-to-one) on the interval $[-\pi/2, \pi/2]$ (see Figure 5.28). On this interval you can define the inverse of the *restricted* sine function to be

$$y = \arcsin x \qquad \text{if and only if} \qquad \sin y = x$$

where $-1 \le x \le 1$ and $-\pi/2 \le \arcsin x \le \pi/2$.

Under suitable restrictions, each of the six trigonometric functions is one-to-one and so has an inverse function, as shown in the following definition.



The sine function is one-to-one on $[-\pi/2, \pi/2]$.
Figure 5.28

NOTE The term "iff" is used to represent the phrase "if and only if."

Definitions of Inverse Trigonometric Functions

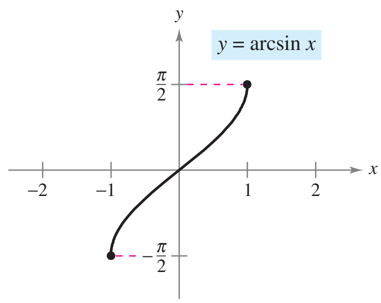
Function	*Domain*	*Range*
$y = \arcsin x$ iff $\sin y = x$	$-1 \le x \le 1$	$-\frac{\pi}{2} \le y \le \frac{\pi}{2}$
$y = \arccos x$ iff $\cos y = x$	$-1 \le x \le 1$	$0 \le y \le \pi$
$y = \arctan x$ iff $\tan y = x$	$-\infty < x < \infty$	$-\frac{\pi}{2} < y < \frac{\pi}{2}$
$y = \operatorname{arccot} x$ iff $\cot y = x$	$-\infty < x < \infty$	$0 < y < \pi$
$y = \operatorname{arcsec} x$ iff $\sec y = x$	$\lvert x\rvert \ge 1$	$0 \le y \le \pi, \quad y \ne \frac{\pi}{2}$
$y = \operatorname{arccsc} x$ iff $\csc y = x$	$\lvert x\rvert \ge 1$	$-\frac{\pi}{2} \le y \le \frac{\pi}{2}, \quad y \ne 0$

NOTE The term "arcsin x" is read as "the arcsine of x" or sometimes "the angle whose sine is x." An alternative notation for the inverse sine function is "$\sin^{-1} x$."

EXPLORATION

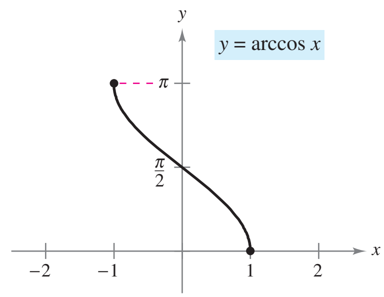
The Inverse Secant Function In the definition above, the inverse secant function is defined by restricting the domain of the secant function to the intervals $\left[0, \frac{\pi}{2}\right) \cup \left(\frac{\pi}{2}, \pi\right]$. Most other texts and reference books agree with this, but some disagree. What other domains might make sense? Explain your reasoning graphically. Most calculators do not have a key for the inverse secant function. How can you use a calculator to evaluate the inverse secant function?

The graphs of the six inverse trigonometric functions are shown in Figure 5.29.



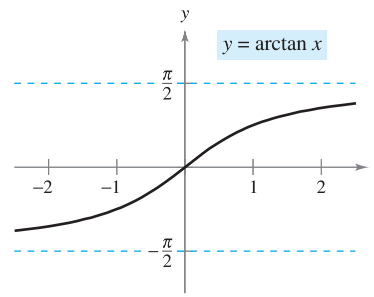
Domain: $[-1, 1]$
Range: $[-\pi/2, \pi/2]$

Animation



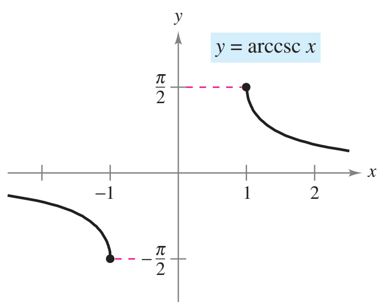
Domain: $[-1, 1]$
Range: $[0, \pi]$

Animation



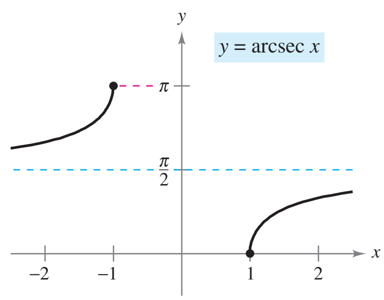
Domain: $(-\infty, \infty)$
Range: $(-\pi/2, \pi/2)$

Animation



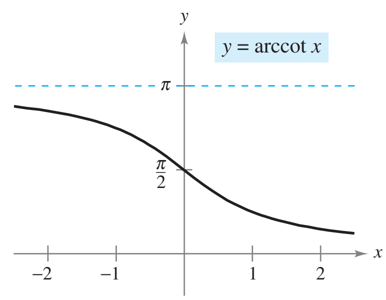
Domain: $(-\infty, -1] \cup [1, \infty)$
Range: $[-\pi/2, 0) \cup (0, \pi/2]$

Animation



Domain: $(-\infty, -1] \cup [1, \infty)$
Range: $[0, \pi/2) \cup (\pi/2, \pi]$

Animation



Domain: $(-\infty, \infty)$
Range: $(0, \pi)$

Animation

Figure 5.29

NOTE When evaluating inverse trigonometric functions, remember that they denote *angles in radian measure*.

EXAMPLE 1 Evaluating Inverse Trigonometric Functions

Evaluate each function.

a. $\arcsin\left(-\dfrac{1}{2}\right)$ **b.** $\arccos 0$ **c.** $\arctan \sqrt{3}$ **d.** $\arcsin(0.3)$

Solution

a. By definition, $y = \arcsin\left(-\frac{1}{2}\right)$ implies that $\sin y = -\frac{1}{2}$. In the interval $[-\pi/2, \pi/2]$, the correct value of y is $-\pi/6$.

$$\arcsin\left(-\frac{1}{2}\right) = -\frac{\pi}{6}$$

b. By definition, $y = \arccos 0$ implies that $\cos y = 0$. In the interval $[0, \pi]$, you have $y = \pi/2$.

$$\arccos 0 = \frac{\pi}{2}$$

c. By definition, $y = \arctan \sqrt{3}$ implies that $\tan y = \sqrt{3}$. In the interval $(-\pi/2, \pi/2)$, you have $y = \pi/3$.

$$\arctan \sqrt{3} = \frac{\pi}{3}$$

d. Using a calculator set in *radian* mode produces

$$\arcsin(0.3) \approx 0.305.$$

Try It Exploration A

EXPLORATION

Graph $y = \arccos(\cos x)$ for $-4\pi \le x \le 4\pi$. Why isn't the graph the same as the graph of $y = x$?

Inverse functions have the properties

$$f(f^{-1}(x)) = x \quad \text{and} \quad f^{-1}(f(x)) = x.$$

When applying these properties to inverse trigonometric functions, remember that the trigonometric functions have inverse functions only in restricted domains. For x-values outside these domains, these two properties do not hold. For example, $\arcsin(\sin \pi)$ is equal to 0, not π.

Properties of Inverse Trigonometric Functions

If $-1 \le x \le 1$ and $-\pi/2 \le y \le \pi/2$, then

$$\sin(\arcsin x) = x \quad \text{and} \quad \arcsin(\sin y) = y.$$

If $-\pi/2 < y < \pi/2$, then

$$\tan(\arctan x) = x \quad \text{and} \quad \arctan(\tan y) = y.$$

If $|x| \ge 1$ and $0 \le y < \pi/2$ or $\pi/2 < y \le \pi$, then

$$\sec(\operatorname{arcsec} x) = x \quad \text{and} \quad \operatorname{arcsec}(\sec y) = y.$$

Similar properties hold for the other inverse trigonometric functions.

Technology

EXAMPLE 2 Solving an Equation

$$\arctan(2x - 3) = \frac{\pi}{4} \qquad \text{Original equation}$$

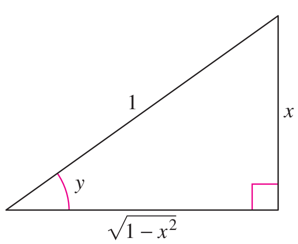
$$\tan[\arctan(2x - 3)] = \tan\frac{\pi}{4} \qquad \text{Take tangent of each side.}$$

$$2x - 3 = 1 \qquad \tan(\arctan x) = x$$

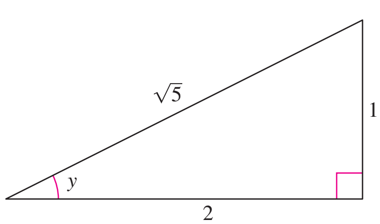
$$x = 2 \qquad \text{Solve for } x.$$

Try It Exploration A Exploration B

Some problems in calculus require that you evaluate expressions such as $\cos(\arcsin x)$, as shown in Example 3.



$y = \arcsin x$
Figure 5.30



$y = \operatorname{arcsec}\dfrac{\sqrt{5}}{2}$
Figure 5.31

EXAMPLE 3 Using Right Triangles

a. Given $y = \arcsin x$, where $0 < y < \pi/2$, find $\cos y$.

b. Given $y = \operatorname{arcsec}(\sqrt{5}/2)$, find $\tan y$.

Solution

a. Because $y = \arcsin x$, you know that $\sin y = x$. This relationship between x and y can be represented by a right triangle, as shown in Figure 5.30.

$$\cos y = \cos(\arcsin x) = \frac{\text{adj.}}{\text{hyp.}} = \sqrt{1 - x^2}$$

(This result is also valid for $-\pi/2 < y < 0$.)

b. Use the right triangle shown in Figure 5.31.

$$\tan y = \tan\left[\operatorname{arcsec}\left(\frac{\sqrt{5}}{2}\right)\right] = \frac{\text{opp.}}{\text{adj.}} = \frac{1}{2}$$

Try It Exploration A

NOTE There is no common agreement on the definition of arcsec x (or arccsc x) for negative values of x. When we defined the range of the arcsecant, we chose to preserve the reciprocal identity

$$\text{arcsec}\, x = \arccos \frac{1}{x}.$$

For example, to evaluate arcsec(-2), you can write

$$\text{arcsec}(-2) = \arccos(-0.5) \approx 2.09.$$

One of the consequences of the definition of the inverse secant function given in this text is that its graph has a positive slope at every x-value in its domain. (See Figure 5.29.) This accounts for the absolute value sign in the formula for the derivative of arcsec x.

Derivatives of Inverse Trigonometric Functions

In Section 5.1 you saw that the derivative of the *transcendental* function $f(x) = \ln x$ is the *algebraic* function $f'(x) = 1/x$. You will now see that the derivatives of the inverse trigonometric functions also are algebraic (even though the inverse trigonometric functions are themselves transcendental).

The following theorem lists the derivatives of the six inverse trigonometric functions. Note that the derivatives of arccos u, arccot u, and arccsc u are the *negatives* of the derivatives of arcsin u, arctan u, and arcsec u, respectively.

THEOREM 5.16 Derivatives of Inverse Trigonometric Functions

Let u be a differentiable function of x.

$$\frac{d}{dx}[\arcsin u] = \frac{u'}{\sqrt{1-u^2}} \qquad \frac{d}{dx}[\arccos u] = \frac{-u'}{\sqrt{1-u^2}}$$

$$\frac{d}{dx}[\arctan u] = \frac{u'}{1+u^2} \qquad \frac{d}{dx}[\text{arccot}\, u] = \frac{-u'}{1+u^2}$$

$$\frac{d}{dx}[\text{arcsec}\, u] = \frac{u'}{|u|\sqrt{u^2-1}} \qquad \frac{d}{dx}[\text{arccsc}\, u] = \frac{-u'}{|u|\sqrt{u^2-1}}$$

Video

To derive these formulas, you can use implicit differentiation. For instance, if $y = \arcsin x$, then $\sin y = x$ and $(\cos y)y' = 1$. (See Exercise 94.)

Technology

TECHNOLOGY If your graphing utility does not have the arcsecant function, you can obtain its graph using

$$f(x) = \text{arcsec}\, x = \arccos \frac{1}{x}.$$

EXAMPLE 4 Differentiating Inverse Trigonometric Functions

a. $\displaystyle \frac{d}{dx}[\arcsin(2x)] = \frac{2}{\sqrt{1-(2x)^2}} = \frac{2}{\sqrt{1-4x^2}}$

b. $\displaystyle \frac{d}{dx}[\arctan(3x)] = \frac{3}{1+(3x)^2} = \frac{3}{1+9x^2}$

c. $\displaystyle \frac{d}{dx}[\arcsin \sqrt{x}] = \frac{(1/2)\,x^{-1/2}}{\sqrt{1-x}} = \frac{1}{2\sqrt{x}\sqrt{1-x}} = \frac{1}{2\sqrt{x-x^2}}$

d. $\displaystyle \frac{d}{dx}[\text{arcsec}\, e^{2x}] = \frac{2e^{2x}}{e^{2x}\sqrt{(e^{2x})^2-1}} = \frac{2e^{2x}}{e^{2x}\sqrt{e^{4x}-1}} = \frac{2}{\sqrt{e^{4x}-1}}$

The absolute value sign is not necessary because $e^{2x} > 0$.

Try It | Exploration A | Exploration B | Video

NOTE From Example 5, you can see one of the benefits of inverse trigonometric functions—they can be used to integrate common algebraic functions. For instance, from the result shown in the example, it follows that

$$\int \sqrt{1-x^2}\,dx = \frac{1}{2}\left(\arcsin x + x\sqrt{1-x^2}\right).$$

EXAMPLE 5 A Derivative That Can Be Simplified

Differentiate $y = \arcsin x + x\sqrt{1-x^2}$.

Solution

$$\begin{aligned} y' &= \frac{1}{\sqrt{1-x^2}} + x\left(\frac{1}{2}\right)(-2x)(1-x^2)^{-1/2} + \sqrt{1-x^2} \\ &= \frac{1}{\sqrt{1-x^2}} - \frac{x^2}{\sqrt{1-x^2}} + \sqrt{1-x^2} \\ &= \sqrt{1-x^2} + \sqrt{1-x^2} \\ &= 2\sqrt{1-x^2} \end{aligned}$$

Try It | Exploration A

The editable graph feature below allows you to edit the graph of a function and its derivative.

Editable Graph

EXAMPLE 6 Analyzing an Inverse Trigonometric Graph

Analyze the graph of $y = (\arctan x)^2$.

Solution From the derivative

$$y' = 2(\arctan x)\left(\frac{1}{1 + x^2}\right) = \frac{2 \arctan x}{1 + x^2}$$

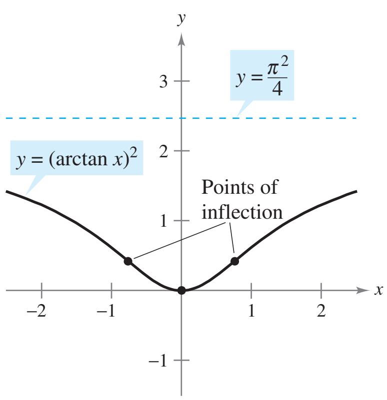
you can see that the only critical number is $x = 0$. By the First Derivative Test, this value corresponds to a relative minimum. From the second derivative

$$y'' = \frac{(1 + x^2)\left(\frac{2}{1 + x^2}\right) - (2 \arctan x)(2x)}{(1 + x^2)^2} = \frac{2(1 - 2x \arctan x)}{(1 + x^2)^2}$$

it follows that points of inflection occur when $2x \arctan x = 1$. Using Newton's Method, these points occur when $x \approx \pm 0.765$. Finally, because

$$\lim_{x \to \pm\infty} (\arctan x)^2 = \frac{\pi^2}{4}$$

it follows that the graph has a horizontal asymptote at $y = \pi^2/4$. The graph is shown in Figure 5.32.



The graph of $y = (\arctan x)^2$ has a horizontal asymptote at $y = \pi^2/4$.
Figure 5.32

Editable Graph

Try It Exploration A

EXAMPLE 7 Maximizing an Angle

A photographer is taking a picture of a four-foot painting hung in an art gallery. The camera lens is 1 foot below the lower edge of the painting, as shown in Figure 5.33. How far should the camera be from the painting to maximize the angle subtended by the camera lens?

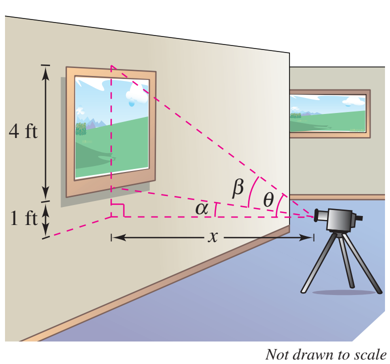
Solution In Figure 5.33, let β be the angle to be maximized.

$$\beta = \theta - \alpha = \operatorname{arccot}\frac{x}{5} - \operatorname{arccot} x$$

Differentiating produces

$$\frac{d\beta}{dx} = \frac{-1/5}{1 + (x^2/25)} - \frac{-1}{1 + x^2} = \frac{-5}{25 + x^2} + \frac{1}{1 + x^2} = \frac{4(5 - x^2)}{(25 + x^2)(1 + x^2)}.$$

Because $d\beta/dx = 0$ when $x = \sqrt{5}$, you can conclude from the First Derivative Test that this distance yields a maximum value of β. So, the distance is $x \approx 2.236$ feet and the angle is $\beta \approx 0.7297$ radian $\approx 41.81°$.



The camera should be 2.236 feet from the painting to maximize the angle β.
Figure 5.33

Try It Open Exploration

GALILEO GALILEI (1564–1642)

Galileo's approach to science departed from the accepted Aristotelian view that nature had describable *qualities*, such as "fluidity" and "potentiality." He chose to describe the physical world in terms of measurable *quantities*, such as time, distance, force, and mass.

MathBio

Review of Basic Differentiation Rules

In the 1600s, Europe was ushered into the scientific age by such great thinkers as Descartes, Galileo, Huygens, Newton, and Kepler. These men believed that nature is governed by basic laws—laws that can, for the most part, be written in terms of mathematical equations. One of the most influential publications of this period—*Dialogue on the Great World Systems*, by Galileo Galilei—has become a classic description of modern scientific thought.

As mathematics has developed during the past few hundred years, a small number of elementary functions has proven sufficient for modeling most* phenomena in physics, chemistry, biology, engineering, economics, and a variety of other fields. An **elementary function** is a function from the following list or one that can be formed as the sum, product, quotient, or composition of functions in the list.

Algebraic Functions	*Transcendental Functions*
Polynomial functions	Logarithmic functions
Rational functions	Exponential functions
Functions involving radicals	Trigonometric functions
	Inverse trigonometric functions

With the differentiation rules introduced so far in the text, you can differentiate *any* elementary function. For convenience, these differentiation rules are summarized below.

Basic Differentiation Rules for Elementary Functions

1. $\dfrac{d}{dx}[cu] = cu'$
2. $\dfrac{d}{dx}[u \pm v] = u' \pm v'$
3. $\dfrac{d}{dx}[uv] = uv' + vu'$
4. $\dfrac{d}{dx}\left[\dfrac{u}{v}\right] = \dfrac{vu' - uv'}{v^2}$
5. $\dfrac{d}{dx}[c] = 0$
6. $\dfrac{d}{dx}[u^n] = nu^{n-1}u'$
7. $\dfrac{d}{dx}[x] = 1$
8. $\dfrac{d}{dx}[|u|] = \dfrac{u}{|u|}(u'), \quad u \neq 0$
9. $\dfrac{d}{dx}[\ln u] = \dfrac{u'}{u}$
10. $\dfrac{d}{dx}[e^u] = e^u\,u'$
11. $\dfrac{d}{dx}[\log_a u] = \dfrac{u'}{(\ln a)u}$
12. $\dfrac{d}{dx}[a^u] = (\ln a)a^u u'$
13. $\dfrac{d}{dx}[\sin u] = (\cos u)u'$
14. $\dfrac{d}{dx}[\cos u] = -(\sin u)u'$
15. $\dfrac{d}{dx}[\tan u] = (\sec^2 u)u'$
16. $\dfrac{d}{dx}[\cot u] = -(\csc^2 u)u'$
17. $\dfrac{d}{dx}[\sec u] = (\sec u \tan u)u'$
18. $\dfrac{d}{dx}[\csc u] = -(\csc u \cot u)u'$
19. $\dfrac{d}{dx}[\arcsin u] = \dfrac{u'}{\sqrt{1 - u^2}}$
20. $\dfrac{d}{dx}[\arccos u] = \dfrac{-u'}{\sqrt{1 - u^2}}$
21. $\dfrac{d}{dx}[\arctan u] = \dfrac{u'}{1 + u^2}$
22. $\dfrac{d}{dx}[\text{arccot } u] = \dfrac{-u'}{1 + u^2}$
23. $\dfrac{d}{dx}[\text{arcsec } u] = \dfrac{u'}{|u|\sqrt{u^2 - 1}}$
24. $\dfrac{d}{dx}[\text{arccsc } u] = \dfrac{-u'}{|u|\sqrt{u^2 - 1}}$

* *Some important functions used in engineering and science (such as Bessel functions and gamma functions) are not elementary functions.*

Section 5.7 Inverse Trigonometric Functions: Integration

- **Integrate functions whose antiderivatives involve inverse trigonometric functions.**
- **Use the method of completing the square to integrate a function.**
- **Review the basic integration rules involving elementary functions.**

Integrals Involving Inverse Trigonometric Functions

The derivatives of the six inverse trigonometric functions fall into three pairs. In each pair, the derivative of one function is the negative of the other. For example,

$$\frac{d}{dx}[\arcsin x] = \frac{1}{\sqrt{1 - x^2}}$$

and

$$\frac{d}{dx}[\arccos x] = -\frac{1}{\sqrt{1 - x^2}}.$$

Technology

When listing the *antiderivative* that corresponds to each of the inverse trigonometric functions, you need to use only one member from each pair. It is conventional to use $\arcsin x$ as the antiderivative of $1/\sqrt{1 - x^2}$, rather than $-\arccos x$. The next theorem gives one antiderivative formula for each of the three pairs. The proofs of these integration rules are left to you (see Exercises 79–81).

FOR FURTHER INFORMATION For a detailed proof of part 2 of Theorem 5.17, see the article "A Direct Proof of the Integral Formula for Arctangent" by Arnold J. Insel in *The College Mathematics Journal*.

MathArticle

THEOREM 5.17 Integrals Involving Inverse Trigonometric Functions

Let u be a differentiable function of x, and let $a > 0$.

1. $\displaystyle\int \frac{du}{\sqrt{a^2 - u^2}} = \arcsin \frac{u}{a} + C$

2. $\displaystyle\int \frac{du}{a^2 + u^2} = \frac{1}{a} \arctan \frac{u}{a} + C$

3. $\displaystyle\int \frac{du}{u\sqrt{u^2 - a^2}} = \frac{1}{a} \operatorname{arcsec} \frac{|u|}{a} + C$

EXAMPLE 1 Integration with Inverse Trigonometric Functions

a. $\displaystyle\int \frac{dx}{\sqrt{4 - x^2}} = \arcsin \frac{x}{2} + C$

b. $\displaystyle\int \frac{dx}{2 + 9x^2} = \frac{1}{3}\int \frac{3\,dx}{(\sqrt{2})^2 + (3x)^2}$ $\qquad u = 3x,\ a = \sqrt{2}$

$$= \frac{1}{3\sqrt{2}} \arctan \frac{3x}{\sqrt{2}} + C$$

c. $\displaystyle\int \frac{dx}{x\sqrt{4x^2 - 9}} = \int \frac{2\,dx}{2x\sqrt{(2x)^2 - 3^2}}$ $\qquad u = 2x,\ a = 3$

$$= \frac{1}{3} \operatorname{arcsec} \frac{|2x|}{3} + C$$

The integrals in Example 1 are fairly straightforward applications of integration formulas. Unfortunately, this is not typical. The integration formulas for inverse trigonometric functions can be disguised in many ways.

Try It Exploration A Video

TECHNOLOGY PITFALL
Computer software that can perform symbolic integration is useful for integrating functions such as the one in Example 2. When using such software, however, you must remember that it can fail to find an antiderivative for two reasons. First, some elementary functions simply do not have antiderivatives that are elementary functions. Second, every symbolic integration utility has limitations—you might have entered a function that the software was not programmed to handle. You should also remember that antiderivatives involving trigonometric functions or logarithmic functions can be written in many different forms. For instance, one symbolic integration utility found the integral in Example 2 to be

$$\int \frac{dx}{\sqrt{e^{2x}-1}} = \arctan \sqrt{e^{2x}-1} + C.$$

Try showing that this antiderivative is equivalent to that obtained in Example 2.

EXAMPLE 2 Integration by Substitution

Find $\displaystyle\int \frac{dx}{\sqrt{e^{2x}-1}}$.

Solution As it stands, this integral doesn't fit any of the three inverse trigonometric formulas. Using the substitution $u = e^x$, however, produces

$$u = e^x \quad \Rightarrow \quad du = e^x\,dx \quad \Rightarrow \quad dx = \frac{du}{e^x} = \frac{du}{u}.$$

With this substitution, you can integrate as follows.

$$\begin{aligned}
\int \frac{dx}{\sqrt{e^{2x}-1}} &= \int \frac{dx}{\sqrt{(e^x)^2-1}} && \text{Write } e^{2x} \text{ as } (e^x)^2.\\
&= \int \frac{du/u}{\sqrt{u^2-1}} && \text{Substitute.}\\
&= \int \frac{du}{u\sqrt{u^2-1}} && \text{Rewrite to fit Arcsecant Rule.}\\
&= \operatorname{arcsec} \frac{|u|}{1} + C && \text{Apply Arcsecant Rule.}\\
&= \operatorname{arcsec} e^x + C && \text{Back-substitute.}
\end{aligned}$$

Try It | Exploration A | Exploration B | Video

EXAMPLE 3 Rewriting as the Sum of Two Quotients

Find $\displaystyle\int \frac{x+2}{\sqrt{4-x^2}}\,dx$.

Solution This integral does not appear to fit any of the basic integration formulas. By splitting the integrand into two parts, however, you can see that the first part can be found with the Power Rule and the second part yields an inverse sine function.

$$\begin{aligned}
\int \frac{x+2}{\sqrt{4-x^2}}\,dx &= \int \frac{x}{\sqrt{4-x^2}}\,dx + \int \frac{2}{\sqrt{4-x^2}}\,dx\\
&= -\frac{1}{2}\int (4-x^2)^{-1/2}(-2x)\,dx + 2\int \frac{1}{\sqrt{4-x^2}}\,dx\\
&= -\frac{1}{2}\left[\frac{(4-x^2)^{1/2}}{1/2}\right] + 2\arcsin\frac{x}{2} + C\\
&= -\sqrt{4-x^2} + 2\arcsin\frac{x}{2} + C
\end{aligned}$$

Try It | Exploration A

Completing the Square

Completing the square helps when quadratic functions are involved in the integrand. For example, the quadratic $x^2 + bx + c$ can be written as the difference of two squares by adding and subtracting $(b/2)^2$.

$$\begin{aligned}
x^2 + bx + c &= x^2 + bx + \left(\frac{b}{2}\right)^2 - \left(\frac{b}{2}\right)^2 + c\\
&= \left(x + \frac{b}{2}\right)^2 - \left(\frac{b}{2}\right)^2 + c
\end{aligned}$$

EXAMPLE 4 Completing the Square

Find $\displaystyle\int \frac{dx}{x^2 - 4x + 7}$.

Solution You can write the denominator as the sum of two squares as shown.

$$\begin{aligned} x^2 - 4x + 7 &= (x^2 - 4x + 4) - 4 + 7 \\ &= (x - 2)^2 + 3 = u^2 + a^2 \end{aligned}$$

Now, in this completed square form, let $u = x - 2$ and $a = \sqrt{3}$.

$$\int \frac{dx}{x^2 - 4x + 7} = \int \frac{dx}{(x-2)^2 + 3} = \frac{1}{\sqrt{3}} \arctan \frac{x-2}{\sqrt{3}} + C$$

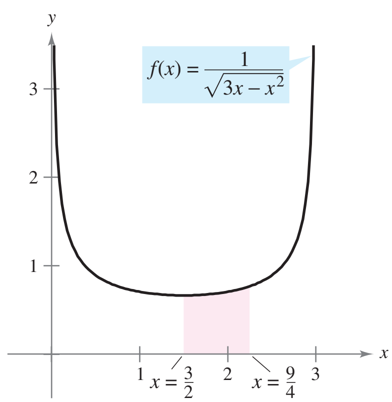
Try It Exploration A Open Exploration

If the leading coefficient is not 1, it helps to factor before completing the square. For instance, you can complete the square of $2x^2 - 8x + 10$ by factoring first.

$$\begin{aligned} 2x^2 - 8x + 10 &= 2(x^2 - 4x + 5) \\ &= 2(x^2 - 4x + 4 - 4 + 5) \\ &= 2[(x - 2)^2 + 1] \end{aligned}$$

To complete the square when the coefficient of x^2 is negative, use the same factoring process shown above. For instance, you can complete the square for $3x - x^2$ as shown.

$$\begin{aligned} 3x - x^2 &= -(x^2 - 3x) \\ &= -\left[x^2 - 3x + \left(\tfrac{3}{2}\right)^2 - \left(\tfrac{3}{2}\right)^2\right] \\ &= \left(\tfrac{3}{2}\right)^2 - \left(x - \tfrac{3}{2}\right)^2 \end{aligned}$$



The area of the region bounded by the graph of f, the x-axis, $x = \frac{3}{2}$, and $x = \frac{9}{4}$ is $\pi/6$.
Figure 5.34

Editable Graph

EXAMPLE 5 Completing the Square (Negative Leading Coefficient)

Find the area of the region bounded by the graph of

$$f(x) = \frac{1}{\sqrt{3x - x^2}}$$

the x-axis, and the lines $x = \frac{3}{2}$ and $x = \frac{9}{4}$.

Solution From Figure 5.34, you can see that the area is given by

$$\text{Area} = \int_{3/2}^{9/4} \frac{1}{\sqrt{3x - x^2}}\, dx.$$

Using the completed square form derived above, you can integrate as shown.

$$\begin{aligned} \int_{3/2}^{9/4} \frac{dx}{\sqrt{3x - x^2}} &= \int_{3/2}^{9/4} \frac{dx}{\sqrt{(3/2)^2 - [x - (3/2)]^2}} \\ &= \arcsin \frac{x - (3/2)}{3/2}\Bigg]_{3/2}^{9/4} \\ &= \arcsin \frac{1}{2} - \arcsin 0 \\ &= \frac{\pi}{6} \\ &\approx 0.524 \end{aligned}$$

Try It Exploration A Exploration B

TECHNOLOGY With definite integrals such as the one given in Example 5, remember that you can resort to a numerical solution. For instance, applying Simpson's Rule (with $n = 12$) to the integral in the example, you obtain

$$\int_{3/2}^{9/4} \frac{1}{\sqrt{3x - x^2}}\, dx \approx 0.523599.$$

This differs from the exact value of the integral ($\pi/6 \approx 0.5235988$) by less than one millionth.

Review of Basic Integration Rules

You have now completed the introduction of the **basic integration rules.** To be efficient at applying these rules, you should have practiced enough so that each rule is committed to memory.

Basic Integration Rules ($a > 0$)

1. $\displaystyle\int kf(u)\,du = k\int f(u)\,du$

2. $\displaystyle\int [f(u) \pm g(u)]\,du = \int f(u)\,du \pm \int g(u)\,du$

3. $\displaystyle\int du = u + C$

4. $\displaystyle\int u^n\,du = \frac{u^{n+1}}{n+1} + C, \quad n \neq -1$

5. $\displaystyle\int \frac{du}{u} = \ln|u| + C$

6. $\displaystyle\int e^u\,du = e^u + C$

7. $\displaystyle\int a^u\,du = \left(\frac{1}{\ln a}\right)a^u + C$

8. $\displaystyle\int \sin u\,du = -\cos u + C$

9. $\displaystyle\int \cos u\,du = \sin u + C$

10. $\displaystyle\int \tan u\,du = -\ln|\cos u| + C$

11. $\displaystyle\int \cot u\,du = \ln|\sin u| + C$

12. $\displaystyle\int \sec u\,du = \ln|\sec u + \tan u| + C$

13. $\displaystyle\int \csc u\,du = -\ln|\csc u + \cot u| + C$

14. $\displaystyle\int \sec^2 u\,du = \tan u + C$

15. $\displaystyle\int \csc^2 u\,du = -\cot u + C$

16. $\displaystyle\int \sec u \tan u\,du = \sec u + C$

17. $\displaystyle\int \csc u \cot u\,du = -\csc u + C$

18. $\displaystyle\int \frac{du}{\sqrt{a^2 - u^2}} = \arcsin\frac{u}{a} + C$

19. $\displaystyle\int \frac{du}{a^2 + u^2} = \frac{1}{a}\arctan\frac{u}{a} + C$

20. $\displaystyle\int \frac{du}{u\sqrt{u^2 - a^2}} = \frac{1}{a}\operatorname{arcsec}\frac{|u|}{a} + C$

You can learn a lot about the nature of integration by comparing this list with the summary of differentiation rules given in the preceding section. For differentiation, you now have rules that allow you to differentiate *any* elementary function. For integration, this is far from true.

The integration rules listed above are primarily those that were happened on when developing differentiation rules. So far, you have not learned any rules or techniques for finding the antiderivative of a general product or quotient, the natural logarithmic function, or the inverse trigonometric functions. More importantly, you cannot apply any of the rules in this list unless you can create the proper du corresponding to the u in the formula. The point is that you need to work more on integration techniques, which you will do in Chapter 8. The next two examples should give you a better feeling for the integration problems that you *can* and *cannot* do with the techniques and rules you now know.

EXAMPLE 6 Comparing Integration Problems

Find as many of the following integrals as you can using the formulas and techniques you have studied so far in the text.

a. $\displaystyle\int \frac{dx}{x\sqrt{x^2-1}}$ **b.** $\displaystyle\int \frac{x\,dx}{\sqrt{x^2-1}}$ **c.** $\displaystyle\int \frac{dx}{\sqrt{x^2-1}}$

Solution

a. You *can* find this integral (it fits the Arcsecant Rule).

$$\int \frac{dx}{x\sqrt{x^2-1}} = \operatorname{arcsec}|x| + C$$

b. You *can* find this integral (it fits the Power Rule).

$$\begin{aligned}\int \frac{x\,dx}{\sqrt{x^2-1}} &= \frac{1}{2}\int (x^2-1)^{-1/2}(2x)\,dx\\ &= \frac{1}{2}\left[\frac{(x^2-1)^{1/2}}{1/2}\right] + C\\ &= \sqrt{x^2-1} + C\end{aligned}$$

c. You *cannot* find this integral using present techniques. (You should scan the list of basic integration rules to verify this conclusion.)

Try It Exploration A

EXAMPLE 7 Comparing Integration Problems

Find as many of the following integrals as you can using the formulas and techniques you have studied so far in the text.

a. $\displaystyle\int \frac{dx}{x\ln x}$ **b.** $\displaystyle\int \frac{\ln x\,dx}{x}$ **c.** $\displaystyle\int \ln x\,dx$

Solution

a. You *can* find this integral (it fits the Log Rule).

$$\begin{aligned}\int \frac{dx}{x\ln x} &= \int \frac{1/x}{\ln x}\,dx\\ &= \ln|\ln x| + C\end{aligned}$$

b. You *can* find this integral (it fits the Power Rule).

$$\begin{aligned}\int \frac{\ln x\,dx}{x} &= \int \left(\frac{1}{x}\right)(\ln x)^1\,dx\\ &= \frac{(\ln x)^2}{2} + C\end{aligned}$$

c. You *cannot* find this integral using present techniques.

Try It Exploration A

NOTE Note in Examples 6 and 7 that the *simplest* functions are the ones that you cannot yet integrate.

Section 5.8 Hyperbolic Functions

- **Develop properties of hyperbolic functions.**
- **Differentiate and integrate hyperbolic functions.**
- **Develop properties of inverse hyperbolic functions.**
- **Differentiate and integrate functions involving inverse hyperbolic functions.**

JOHANN HEINRICH LAMBERT (1728–1777)
The first person to publish a comprehensive study on hyperbolic functions was Johann Heinrich Lambert, a Swiss-German mathematician and colleague of Euler.

MathBio

Hyperbolic Functions

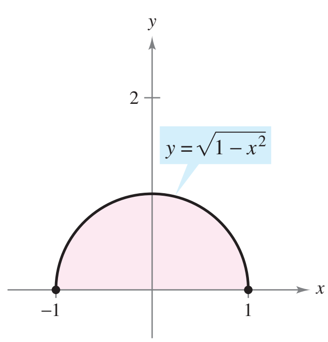
In this section you will look briefly at a special class of exponential functions called **hyperbolic functions.** The name *hyperbolic function* arose from comparison of the area of a semicircular region, as shown in Figure 5.35, with the area of a region under a hyperbola, as shown in Figure 5.36. The integral for the semicircular region involves an inverse trigonometric (circular) function:

$$\int_{-1}^{1} \sqrt{1-x^2}\,dx = \frac{1}{2}\Big[x\sqrt{1-x^2} + \arcsin x\Big]_{-1}^{1} = \frac{\pi}{2} \approx 1.571.$$

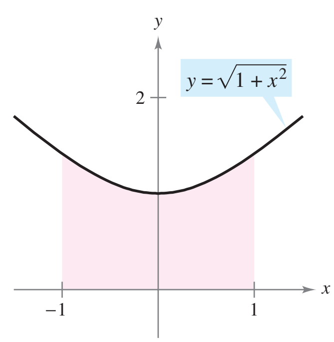
The integral for the hyperbolic region involves an inverse hyperbolic function:

$$\int_{-1}^{1} \sqrt{1+x^2}\,dx = \frac{1}{2}\Big[x\sqrt{1+x^2} + \sinh^{-1} x\Big]_{-1}^{1} \approx 2.296.$$

This is only one of many ways in which the hyperbolic functions are similar to the trigonometric functions.



Circle: $x^2 + y^2 = 1$
Figure 5.35



Hyperbola: $-x^2 + y^2 = 1$
Figure 5.36

FOR FURTHER INFORMATION For more information on the development of hyperbolic functions, see the article "An Introduction to Hyperbolic Functions in Elementary Calculus" by Jerome Rosenthal in *Mathematics Teacher*.

MathArticle

Definitions of the Hyperbolic Functions

$$\sinh x = \frac{e^x - e^{-x}}{2} \qquad \operatorname{csch} x = \frac{1}{\sinh x}, \quad x \neq 0$$

$$\cosh x = \frac{e^x + e^{-x}}{2} \qquad \operatorname{sech} x = \frac{1}{\cosh x}$$

$$\tanh x = \frac{\sinh x}{\cosh x} \qquad \coth x = \frac{1}{\tanh x}, \quad x \neq 0$$

Video

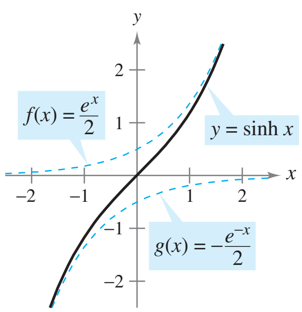
NOTE sinh x is read as "the hyperbolic sine of x," cosh x as "the hyperbolic cosine of x," and so on.

The graphs of the six hyperbolic functions and their domains and ranges are shown in Figure 5.37. Note that the graph of $\sinh x$ can be obtained by *addition of ordinates* using the exponential functions $f(x) = \frac{1}{2}e^x$ and $g(x) = -\frac{1}{2}e^{-x}$. View the animations to see this.

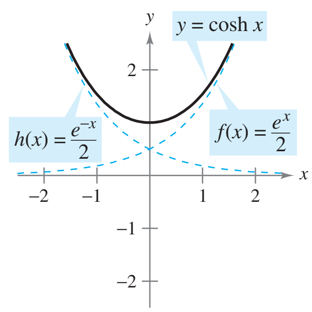
$$f(x) = \sinh x = \frac{1}{2}e^x - \frac{1}{2}e^{-x} \qquad f(x) = \cosh x = \frac{1}{2}e^x - \frac{1}{2}e^{-x}$$

Animation Animation

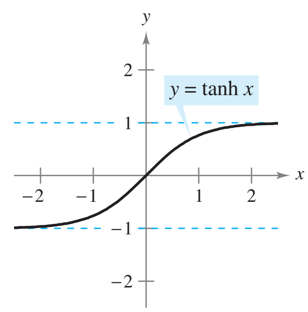
Likewise, the graph of $\cosh x$ can be obtained by *addition of ordinates* using the exponential functions $f(x) = \frac{1}{2}e^x$ and $h(x) = \frac{1}{2}e^{-x}$.



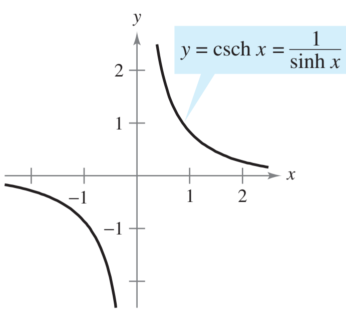
Domain: $(-\infty, \infty)$
Range: $(-\infty, \infty)$



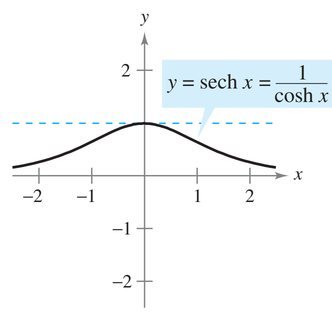
Domain: $(-\infty, \infty)$
Range: $[1, \infty)$



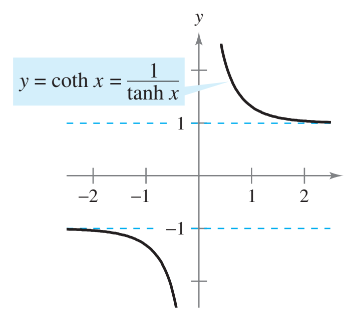
Domain: $(-\infty, \infty)$
Range: $(-1, 1)$



Domain: $(-\infty, 0) \cup (0, \infty)$
Range: $(-\infty, 0) \cup (0, \infty)$



Domain: $(-\infty, \infty)$
Range: $(0, 1]$



Domain: $(-\infty, 0) \cup (0, \infty)$
Range: $(-\infty, -1) \cup (1, \infty)$

Figure 5.37

Many of the trigonometric identities have corresponding *hyperbolic identities*. For instance,

$$\begin{aligned}
\cosh^2 x - \sinh^2 x &= \left(\frac{e^x + e^{-x}}{2}\right)^2 - \left(\frac{e^x - e^{-x}}{2}\right)^2 \\
&= \frac{e^{2x} + 2 + e^{-2x}}{4} - \frac{e^{2x} - 2 + e^{-2x}}{4} \\
&= \frac{4}{4} \\
&= 1
\end{aligned}$$

and

$$\begin{aligned}
2 \sinh x \cosh x &= 2\left(\frac{e^x - e^{-x}}{2}\right)\left(\frac{e^x + e^{-x}}{2}\right) \\
&= \frac{e^{2x} - e^{-2x}}{2}
\end{aligned}$$

Hyperbolic Identities

$$\cosh^2 x - \sinh^2 x = 1 \qquad \sinh(x + y) = \sinh x \cosh y + \cosh x \sinh y$$
$$\tanh^2 x + \operatorname{sech}^2 x = 1 \qquad \sinh(x - y) = \sinh x \cosh y - \cosh x \sinh y$$
$$\coth^2 x - \operatorname{csch}^2 x = 1 \qquad \cosh(x + y) = \cosh x \cosh y + \sinh x \sinh y$$
$$\cosh(x - y) = \cosh x \cosh y - \sinh x \sinh y$$
$$\sinh^2 x = \frac{-1 + \cosh 2x}{2} \qquad \cosh^2 x = \frac{1 + \cosh 2x}{2}$$
$$\sinh 2x = 2 \sinh x \cosh x \qquad \cosh 2x = \cosh^2 x + \sinh^2 x$$

Differentiation and Integration of Hyperbolic Functions

Because the hyperbolic functions are written in terms of e^x and e^{-x}, you can easily derive rules for their derivatives. The following theorem lists these derivatives with the corresponding integration rules.

THEOREM 5.18 Derivatives and Integrals of Hyperbolic Functions

Let u be a differentiable function of x.

$$\frac{d}{dx}[\sinh u] = (\cosh u)u' \qquad \int \cosh u\, du = \sinh u + C$$
$$\frac{d}{dx}[\cosh u] = (\sinh u)u' \qquad \int \sinh u\, du = \cosh u + C$$
$$\frac{d}{dx}[\tanh u] = (\operatorname{sech}^2 u)u' \qquad \int \operatorname{sech}^2 u\, du = \tanh u + C$$
$$\frac{d}{dx}[\coth u] = -(\operatorname{csch}^2 u)u' \qquad \int \operatorname{csch}^2 u\, du = -\coth u + C$$
$$\frac{d}{dx}[\operatorname{sech} u] = -(\operatorname{sech} u \tanh u)u' \qquad \int \operatorname{sech} u \tanh u\, du = -\operatorname{sech} u + C$$
$$\frac{d}{dx}[\operatorname{csch} u] = -(\operatorname{csch} u \coth u)u' \qquad \int \operatorname{csch} u \coth u\, du = -\operatorname{csch} u + C$$

Proof

$$\begin{aligned}\frac{d}{dx}[\sinh x] &= \frac{d}{dx}\left[\frac{e^x - e^{-x}}{2}\right]\\ &= \frac{e^x + e^{-x}}{2} = \cosh x\\ \frac{d}{dx}[\tanh x] &= \frac{d}{dx}\left[\frac{\sinh x}{\cosh x}\right]\\ &= \frac{\cosh x(\cosh x) - \sinh x(\sinh x)}{\cosh^2 x}\\ &= \frac{1}{\cosh^2 x}\\ &= \operatorname{sech}^2 x\end{aligned}$$

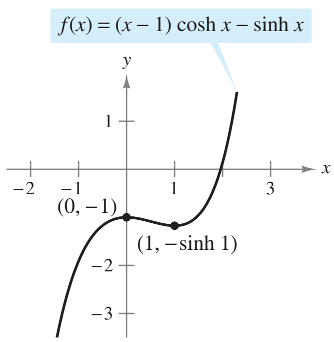
In Exercises 98 and 102, you are asked to prove some of the other differentiation rules.

EXAMPLE 1 Differentiation of Hyperbolic Functions

a. $\dfrac{d}{dx}[\sinh(x^2-3)] = 2x\cosh(x^2-3)$ **b.** $\dfrac{d}{dx}[\ln(\cosh x)] = \dfrac{\sinh x}{\cosh x} = \tanh x$

c. $\dfrac{d}{dx}[x\sinh x - \cosh x] = x\cosh x + \sinh x - \sinh x = x\cosh x$

Try It | Exploration A | Exploration B | Video



$f''(0) < 0$, so $(0, -1)$ is a relative maximum. $f''(1) > 0$, so $(1, -\sinh 1)$ is a relative minimum.
Figure 5.38

EXAMPLE 2 Finding Relative Extrema

Find the relative extrema of $f(x) = (x-1)\cosh x - \sinh x$.

Solution Begin by setting the first derivative of f equal to 0.

$$\begin{aligned} f'(x) = (x-1)\sinh x + \cosh x - \cosh x &= 0 \\ (x-1)\sinh x &= 0 \end{aligned}$$

So, the critical numbers are $x = 1$ and $x = 0$. Using the Second Derivative Test, you can verify that the point $(0, -1)$ yields a relative maximum and the point $(1, -\sinh 1)$ yields a relative minimum, as shown in Figure 5.38. Try using a graphing utility to confirm this result. If your graphing utility does not have hyperbolic functions, you can use exponential functions as follows.

$$\begin{aligned} f(x) &= (x-1)\left(\tfrac{1}{2}\right)(e^x + e^{-x}) - \tfrac{1}{2}(e^x - e^{-x}) \\ &= \tfrac{1}{2}(xe^x + xe^{-x} - e^x - e^{-x} - e^x + e^{-x}) \\ &= \tfrac{1}{2}(xe^x + xe^{-x} - 2e^x) \end{aligned}$$

Try It | Exploration A

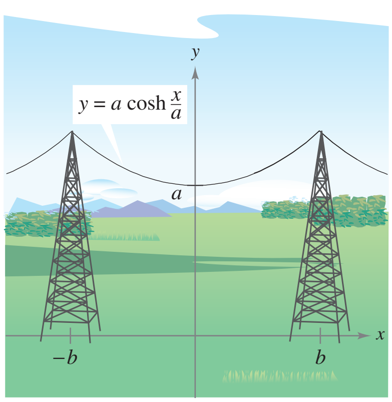
When a uniform flexible cable, such as a telephone wire, is suspended from two points, it takes the shape of a *catenary*, as discussed in Example 3.

EXAMPLE 3 Hanging Power Cables

Power cables are suspended between two towers, forming the catenary shown in Figure 5.39. The equation for this catenary is

$$y = a\cosh\frac{x}{a}.$$

The distance between the two towers is $2b$. Find the slope of the catenary at the point where the cable meets the right-hand tower.



Catenary
Figure 5.39

Solution Differentiating produces

$$y' = a\left(\frac{1}{a}\right)\sinh\frac{x}{a} = \sinh\frac{x}{a}.$$

At the point $(b, a\cosh(b/a))$, the slope (from the left) is given by $m = \sinh\dfrac{b}{a}$.

Try It | Exploration A | Open Exploration

FOR FURTHER INFORMATION In Example 3, the cable is a catenary between two supports at the same height. To learn about the shape of a cable hanging between supports of different heights, see the article "Reexamining the Catenary" by Paul Cella in *The College Mathematics Journal*.

EXAMPLE 4 Integrating a Hyperbolic Function

Find $\displaystyle\int \cosh 2x \sinh^2 2x\, dx$.

Solution

$$\int \cosh 2x \sinh^2 2x\, dx = \frac{1}{2}\int (\sinh 2x)^2(2\cosh 2x)\, dx \qquad u = \sinh 2x$$
$$= \frac{1}{2}\left[\frac{(\sinh 2x)^3}{3}\right] + C$$
$$= \frac{\sinh^3 2x}{6} + C$$

Try It   Exploration A

Inverse Hyperbolic Functions

Unlike trigonometric functions, hyperbolic functions are not periodic. In fact, by looking back at Figure 5.37, you can see that four of the six hyperbolic functions are actually one-to-one (the hyperbolic sine, tangent, cosecant, and cotangent). So, you can apply Theorem 5.7 to conclude that these four functions have inverse functions. The other two (the hyperbolic cosine and secant) are one-to-one if their domains are restricted to the positive real numbers, and for this restricted domain they also have inverse functions. Because the hyperbolic functions are defined in terms of exponential functions, it is not surprising to find that the inverse hyperbolic functions can be written in terms of logarithmic functions, as shown in Theorem 5.19.

THEOREM 5.19 Inverse Hyperbolic Functions

Function	*Domain*
$\sinh^{-1} x = \ln\left(x + \sqrt{x^2+1}\right)$	$(-\infty, \infty)$
$\cosh^{-1} x = \ln\left(x + \sqrt{x^2-1}\right)$	$[1, \infty)$
$\tanh^{-1} x = \dfrac{1}{2}\ln\dfrac{1+x}{1-x}$	$(-1, 1)$
$\coth^{-1} x = \dfrac{1}{2}\ln\dfrac{x+1}{x-1}$	$(-\infty, -1) \cup (1, \infty)$
$\operatorname{sech}^{-1} x = \ln\dfrac{1+\sqrt{1-x^2}}{x}$	$(0, 1]$
$\operatorname{csch}^{-1} x = \ln\left(\dfrac{1}{x} + \dfrac{\sqrt{1+x^2}}{\lvert x\rvert}\right)$	$(-\infty, 0) \cup (0, \infty)$

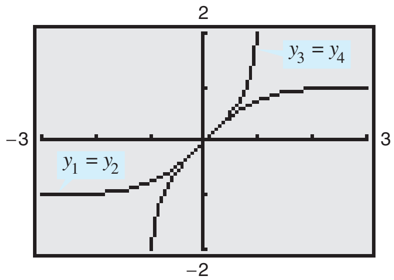
Proof The proof of this theorem is a straightforward application of the properties of the exponential and logarithmic functions. For example, if

$$f(x) = \sinh x = \frac{e^x - e^{-x}}{2}$$

and

$$g(x) = \ln\left(x + \sqrt{x^2+1}\right)$$

you can show that $f(g(x)) = x$ and $g(f(x)) = x$, which implies that g is the inverse function of f.



Graphs of the hyperbolic tangent function and the inverse hyperbolic tangent function
Figure 5.40

TECHNOLOGY You can use a graphing utility to confirm graphically the results of Theorem 5.19. For instance, graph the following functions.

$$y_1 = \tanh x \qquad \text{Hyperbolic tangent}$$

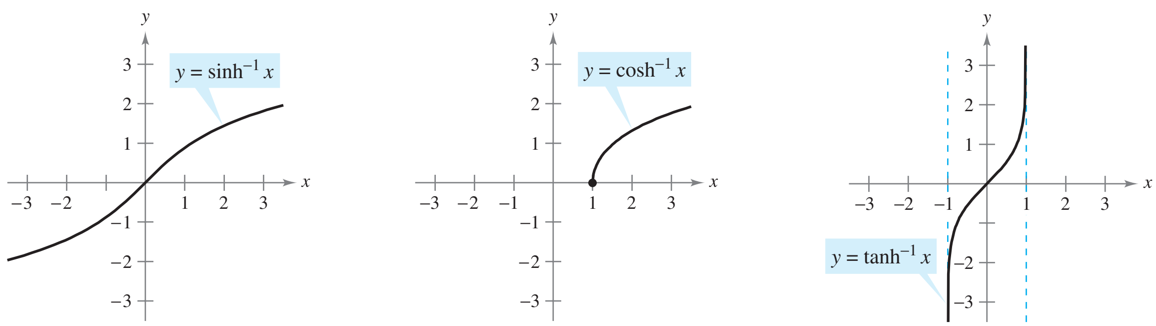
$$y_2 = \frac{e^x - e^{-x}}{e^x + e^{-x}} \qquad \text{Definition of hyperbolic tangent}$$

$$y_3 = \tanh^{-1} x \qquad \text{Inverse hyperbolic tangent}$$

$$y_4 = \frac{1}{2} \ln \frac{1 + x}{1 - x} \qquad \text{Definition of inverse hyperbolic tangent}$$

The resulting display is shown in Figure 5.40. As you watch the graphs being traced out, notice that $y_1 = y_2$ and $y_3 = y_4$. Also notice that the graph of y_1 is the reflection of the graph of y_3 in the line $y = x$.

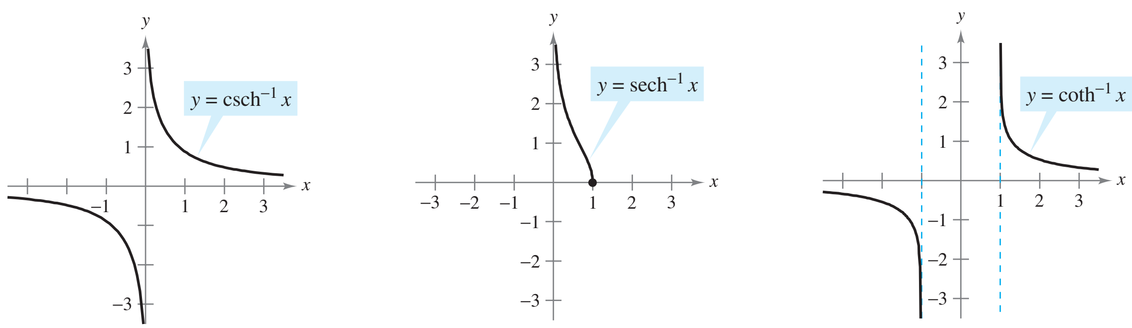
The graphs of the inverse hyperbolic functions are shown in Figure 5.41.



Domain: $(-\infty, \infty)$
Range: $(-\infty, \infty)$

Domain: $[1, \infty)$
Range: $[0, \infty)$

Domain: $(-1, 1)$
Range: $(-\infty, \infty)$



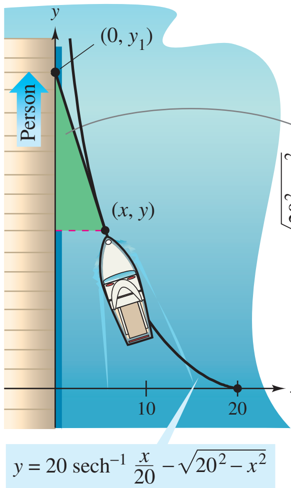
Domain: $(-\infty, 0) \cup (0, \infty)$
Range: $(-\infty, 0) \cup (0, \infty)$

Domain: $(0, 1]$
Range: $[0, \infty)$

Domain: $(-\infty, -1) \cup (1, \infty)$
Range: $(-\infty, 0) \cup (0, \infty)$

Figure 5.41

The inverse hyperbolic secant can be used to define a curve called a *tractrix* or *pursuit curve*, as discussed in Example 5.



A person must walk 41.27 feet to bring the boat 5 feet from the dock.
Figure 5.42

EXAMPLE 5 A Tractrix

A person is holding a rope that is tied to a boat, as shown in Figure 5.42. As the person walks along the dock, the boat travels along a **tractrix,** given by the equation

$$y = a \operatorname{sech}^{-1} \frac{x}{a} - \sqrt{a^2 - x^2}$$

where a is the length of the rope. If $a = 20$ feet, find the distance the person must walk to bring the boat 5 feet from the dock.

Solution In Figure 5.42, notice that the distance the person has walked is given by

$$\begin{aligned} y_1 = y + \sqrt{20^2 - x^2} &= \left(20 \operatorname{sech}^{-1} \frac{x}{20} - \sqrt{20^2 - x^2}\right) + \sqrt{20^2 - x^2} \\ &= 20 \operatorname{sech}^{-1} \frac{x}{20}. \end{aligned}$$

When $x = 5$, this distance is

$$\begin{aligned} y_1 = 20 \operatorname{sech}^{-1} \frac{5}{20} &= 20 \ln \frac{1 + \sqrt{1 - (1/4)^2}}{1/4} \\ &= 20 \ln\left(4 + \sqrt{15}\right) \\ &\approx 41.27 \text{ feet.} \end{aligned}$$

Try It Exploration A

Differentiation and Integration of Inverse Hyperbolic Functions

The derivatives of the inverse hyperbolic functions, which resemble the derivatives of the inverse trigonometric functions, are listed in Theorem 5.20 with the corresponding integration formulas (in logarithmic form). You can verify each of these formulas by applying the logarithmic definitions of the inverse hyperbolic functions. (See Exercises 99–101.)

THEOREM 5.20 Differentiation and Integration Involving Inverse Hyperbolic Functions

Let u be a differentiable function of x.

$$\frac{d}{dx}[\sinh^{-1} u] = \frac{u'}{\sqrt{u^2 + 1}} \qquad \frac{d}{dx}[\cosh^{-1} u] = \frac{u'}{\sqrt{u^2 - 1}}$$

$$\frac{d}{dx}[\tanh^{-1} u] = \frac{u'}{1 - u^2} \qquad \frac{d}{dx}[\coth^{-1} u] = \frac{u'}{1 - u^2}$$

$$\frac{d}{dx}[\operatorname{sech}^{-1} u] = \frac{-u'}{u\sqrt{1 - u^2}} \qquad \frac{d}{dx}[\operatorname{csch}^{-1} u] = \frac{-u'}{|u|\sqrt{1 + u^2}}$$

$$\int \frac{du}{\sqrt{u^2 \pm a^2}} = \ln\left(u + \sqrt{u^2 \pm a^2}\right) + C$$

$$\int \frac{du}{a^2 - u^2} = \frac{1}{2a} \ln\left|\frac{a + u}{a - u}\right| + C$$

$$\int \frac{du}{u\sqrt{a^2 \pm u^2}} = -\frac{1}{a} \ln \frac{a + \sqrt{a^2 \pm u^2}}{|u|} + C$$

EXAMPLE 6 More About a Tractrix

For the tractrix given in Example 5, show that the boat is always pointing toward the person.

Solution For a point (x, y) on a tractrix, the slope of the graph gives the direction of the boat, as shown in Figure 5.42.

$$\begin{aligned} y' &= \frac{d}{dx}\left[20 \operatorname{sech}^{-1}\frac{x}{20} - \sqrt{20^2 - x^2}\right] \\ &= -20\left(\frac{1}{20}\right)\left[\frac{1}{(x/20)\sqrt{1-(x/20)^2}}\right] - \left(\frac{1}{2}\right)\left(\frac{-2x}{\sqrt{20^2 - x^2}}\right) \\ &= \frac{-20^2}{x\sqrt{20^2 - x^2}} + \frac{x}{\sqrt{20^2 - x^2}} \\ &= -\frac{\sqrt{20^2 - x^2}}{x} \end{aligned}$$

However, from Figure 5.42, you can see that the slope of the line segment connecting the point $(0, y_1)$ with the point (x, y) is also

$$m = -\frac{\sqrt{20^2 - x^2}}{x}.$$

So, the boat is always pointing toward the person. (It is because of this property that a tractrix is called a *pursuit curve*.)

Try It **Exploration A**

EXAMPLE 7 Integration Using Inverse Hyperbolic Functions

Find $\displaystyle\int \frac{dx}{x\sqrt{4 - 9x^2}}$.

Solution Let $a = 2$ and $u = 3x$.

$$\begin{aligned} \int \frac{dx}{x\sqrt{4-9x^2}} &= \int \frac{3\,dx}{(3x)\sqrt{4-9x^2}} && \int \frac{du}{u\sqrt{a^2-u^2}} \\ &= -\frac{1}{2}\ln\frac{2+\sqrt{4-9x^2}}{|3x|} + C && -\frac{1}{a}\ln\frac{a+\sqrt{a^2-u^2}}{|u|} + C \end{aligned}$$

Try It **Exploration A**

EXAMPLE 8 Integration Using Inverse Hyperbolic Functions

Find $\displaystyle\int \frac{dx}{5 - 4x^2}$.

Solution Let $a = \sqrt{5}$ and $u = 2x$.

$$\begin{aligned} \int \frac{dx}{5-4x^2} &= \frac{1}{2}\int \frac{2\,dx}{(\sqrt{5})^2 - (2x)^2} && \int \frac{du}{a^2-u^2} \\ &= \frac{1}{2}\left(\frac{1}{2\sqrt{5}}\ln\left|\frac{\sqrt{5}+2x}{\sqrt{5}-2x}\right|\right) + C && \frac{1}{2a}\ln\left|\frac{a+u}{a-u}\right| + C \\ &= \frac{1}{4\sqrt{5}}\ln\left|\frac{\sqrt{5}+2x}{\sqrt{5}-2x}\right| + C \end{aligned}$$

Try It **Exploration A**

Section 6.1 Slope Fields and Euler's Method

- **Use initial conditions to find particular solutions of differential equations.**
- **Use slope fields to approximate solutions of differential equations.**
- **Use Euler's Method to approximate solutions of differential equations.**

General and Particular Solutions

In this text, you will learn that physical phenomena can be described by differential equations. In Section 6.2, you will see that problems involving radioactive decay, population growth, and Newton's Law of Cooling can be formulated in terms of differential equations.

A function $y = f(x)$ is called a **solution** of a differential equation if the equation is satisfied when y and its derivatives are replaced by $f(x)$ and its derivatives. For example, differentiation and substitution would show that $y = e^{-2x}$ is a solution of the differential equation $y' + 2y = 0$. It can be shown that every solution of this differential equation is of the form

$$y = Ce^{-2x} \qquad \text{General solution of } y' + 2y = 0$$

where C is any real number. This solution is called the **general solution.** Some differential equations have **singular solutions** that cannot be written as special cases of the general solution. However, such solutions are not considered in this text. The **order** of a differential equation is determined by the highest-order derivative in the equation. For instance, $y' = 4y$ is a first-order differential equation.

NOTE First-order linear differential equations are discussed in Section 6.4.

In Section 4.1, Example 8, you saw that the second-order differential equation $s''(t) = -32$ has the general solution

$$s(t) = -16t^2 + C_1t + C_2 \qquad \text{General solution of } s''(t) = -32$$

which contains two arbitrary constants. It can be shown that a differential equation of order n has a general solution with n arbitrary constants.

EXAMPLE 1 Verifying Solutions

Determine whether the function is a solution of the differential equation $y'' - y = 0$.

a. $y = \sin x$ **b.** $y = 4e^{-x}$ **c.** $y = Ce^x$

Solution

a. Because $y = \sin x$, $y' = \cos x$, and $y'' = -\sin x$, it follows that

$$y'' - y = -\sin x - \sin x = -2 \sin x \neq 0.$$

So, $y = \sin x$ is *not* a solution.

b. Because $y = 4e^{-x}$, $y' = -4e^{-x}$, and $y'' = 4e^{-x}$, it follows that

$$y'' - y = 4e^{-x} - 4e^{-x} = 0.$$

So, $y = 4e^{-x}$ is a solution.

c. Because $y = Ce^x$, $y' = Ce^x$, and $y'' = Ce^x$, it follows that

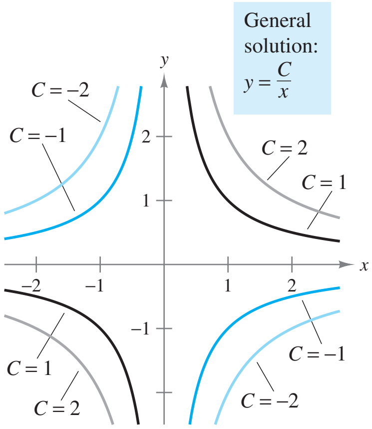
$$y'' - y = Ce^x - Ce^x = 0.$$

So, $y = Ce^x$ is a solution for any value of C.

Try It Exploration A

The editable graph feature below allows you to edit the graph of a function.

Editable Graph



Solution curves for $xy' + y = 0$
Figure 6.1

Geometrically, the general solution of a first-order differential equation represents a family of curves known as **solution curves,** one for each value assigned to the arbitrary constant. For instance, you can verify that every function of the form

$$y = \frac{C}{x} \qquad \text{General solution of } xy' + y = 0$$

is a solution of the differential equation $xy' + y = 0$. Figure 6.1 shows four of the solution curves corresponding to different values of C.

As discussed in Section 4.1, **particular solutions** of a differential equation are obtained from **initial conditions** that give the value of the dependent variable or one of its derivatives for a particular value of the independent variable. The term "initial condition" stems from the fact that, often in problems involving time, the value of the dependent variable or one of its derivatives is known at the *initial* time $t = 0$. For instance, the second-order differential equation $s''(t) = -32$ having the general solution

$$s(t) = -16t^2 + C_1t + C_2 \qquad \text{General solution of } s''(t) = -32$$

might have the following initial conditions.

$$s(0) = 80, \quad s'(0) = 64 \qquad \text{Initial conditions}$$

In this case, the initial conditions yield the particular solution

$$s(t) = -16t^2 + 64t + 80. \qquad \text{Particular solution}$$

EXAMPLE 2 Finding a Particular Solution

For the differential equation $xy' - 3y = 0$, verify that $y = Cx^3$ is a solution, and find the particular solution determined by the initial condition $y = 2$ when $x = -3$.

Solution You know that $y = Cx^3$ is a solution because $y' = 3Cx^2$ and

$$\begin{aligned} xy' - 3y &= x(3Cx^2) - 3(Cx^3) \\ &= 0. \end{aligned}$$

Furthermore, the initial condition $y = 2$ when $x = -3$ yields

$$\begin{aligned} y &= Cx^3 && \text{General solution} \\ 2 &= C(-3)^3 && \text{Substitute initial condition.} \\ -\frac{2}{27} &= C && \text{Solve for } C. \end{aligned}$$

and you can conclude that the particular solution is

$$y = -\frac{2x^3}{27}. \qquad \text{Particular solution}$$

Try checking this solution by substituting for y and y' in the original differential equation.

Try It | Open Exploration

The editable graph feature below allows you to edit the graph of a function.

Editable Graph

NOTE To determine a particular solution, the number of initial conditions must match the number of constants in the general solution.

Slope Fields

Solving a differential equation analytically can be difficult or even impossible. However, there is a graphical approach you can use to learn a lot about the solution of a differential equation. Consider a differential equation of the form

$$y' = F(x, y). \qquad \text{Differential equation}$$

At each point (x, y) in the xy-plane where F is defined, the differential equation determines the slope $y' = F(x, y)$ of the solution at that point. If you draw a short line segment with slope $F(x, y)$ at selected points (x, y) in the domain of F, then these line segments form a **slope field,** or a *direction field* for the differential equation $y' = F(x, y)$. Each line segment has the same slope as the solution curve through that point. A slope field shows the general shape of all the solutions.

EXAMPLE 3 Sketching a Slope Field

Sketch a slope field for the differential equation $y' = x - y$ for the points $(-1, 1)$, $(0, 1)$, and $(1, 1)$.

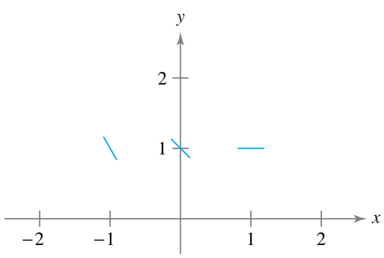


Figure 6.2

Solution

The slope of the solution curve at any point (x, y) is $F(x, y) = x - y$. So, the slope at $(-1, 1)$ is $y' = -1 - 1 = -2$, the slope at $(0, 1)$ is $y' = 0 - 1 = -1$, and the slope at $(1, 1)$ is $y' = 1 - 1 = 0$. Draw short line segments at the three points with their respective slopes, as shown in Figure 6.2.

Try It | Exploration A

EXAMPLE 4 Identifying Slope Fields for Differential Equations

Match each slope field with its differential equation.

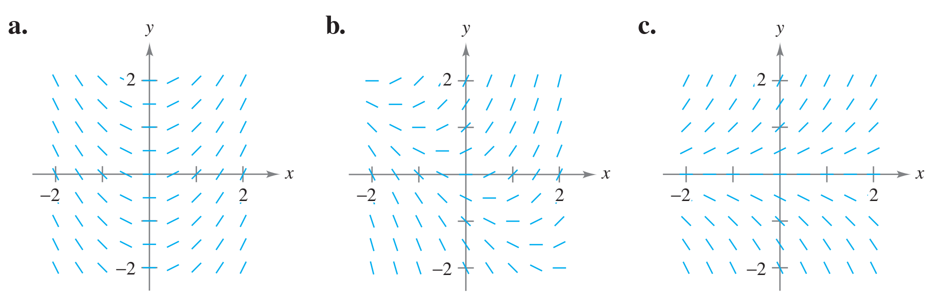


Figure 6.3

i. $y' = x + y$ **ii.** $y' = x$ **iii.** $y' = y$

Solution

a. From Figure 6.3(a), you can see that the slope at any point along the y-axis is 0. The only equation that satisfies this condition is $y' = x$. So, the graph matches (ii).

b. From Figure 6.3(b), you can see that the slope at the point $(1, -1)$ is 0. The only equation that satisfies this condition is $y' = x + y$. So, the graph matches (i).

c. From Figure 6.3(c), you can see that the slope at any point along the x-axis is 0. The only equation that satisfies this condition is $y' = y$. So, the graph matches (iii).

Try It | Exploration A

A solution curve of a differential equation $y' = F(x, y)$ is simply a curve in the xy-plane whose tangent line at each point (x, y) has slope equal to $F(x, y)$. This is illustrated in Example 5.

EXAMPLE 5 Sketching a Solution Using a Slope Field

Sketch a slope field for the differential equation

$$y' = 2x + y.$$

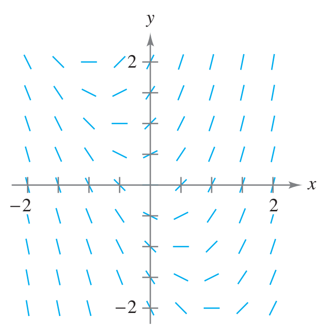
Use the slope field to sketch the solution that passes through the point $(1, 1)$.

Solution

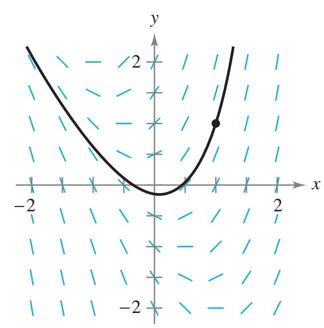
Make a table showing the slopes at several points. The table shown is a small sample. The slopes at many other points should be calculated to get a representative slope field.

x	-2	-2	-1	-1	0	0	1	1	2	2
y	-1	1	-1	1	-1	1	-1	1	-1	1
$y' = 2x + y$	-5	-3	-3	-1	-1	1	1	3	3	5

Next draw line segments at the points with their respective slopes, as shown in Figure 6.4.



Slope field for $y' = 2x + y$
Figure 6.4



Particular solution for $y' = 2x + y$ passing through $(1, 1)$
Figure 6.5

After the slope field is drawn, start at the initial point $(1, 1)$ and move to the right in the direction of the line segment. Continue to draw the solution curve so that it moves parallel to the nearby line segments. Do the same to the left of $(1, 1)$. The resulting solution is shown in Figure 6.5.

Try It **Exploration A**

From Example 5, note that the slope field shows that y' increases to infinity as x increases.

NOTE Drawing a slope field by hand is tedious. In practice, slope fields are usually drawn using a graphing utility.

Euler's Method

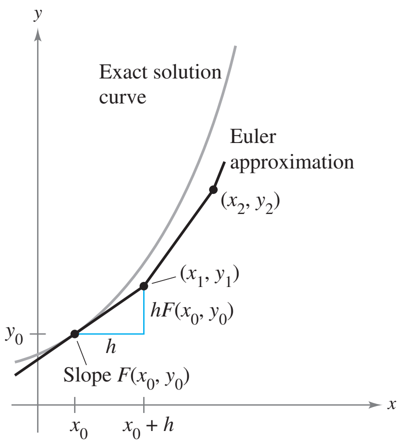


Figure 6.6

Euler's Method is a numerical approach to approximating the particular solution of the differential equation

$$y' = F(x, y)$$

that passes through the point (x_0, y_0). From the given information, you know that the graph of the solution passes through the point (x_0, y_0) and has a slope of $F(x_0, y_0)$ at this point. This gives you a "starting point" for approximating the solution.

From this starting point, you can proceed in the direction indicated by the slope. Using a small step h, move along the tangent line until you arrive at the point (x_1, y_1), where

$$x_1 = x_0 + h \quad \text{and} \quad y_1 = y_0 + hF(x_0, y_0)$$

as shown in Figure 6.6. If you think of (x_1, y_1) as a new starting point, you can repeat the process to obtain a second point (x_2, y_2). The values of x_i and y_i are as follows.

$$\begin{aligned}
x_1 &= x_0 + h & y_1 &= y_0 + hF(x_0, y_0) \\
x_2 &= x_1 + h & y_2 &= y_1 + hF(x_1, y_1) \\
&\vdots & &\vdots \\
x_n &= x_{n-1} + h & y_n &= y_{n-1} + hF(x_{n-1}, y_{n-1})
\end{aligned}$$

NOTE You can obtain better approximations of the exact solution by choosing smaller and smaller step sizes.

EXAMPLE 6 Approximating a Solution Using Euler's Method

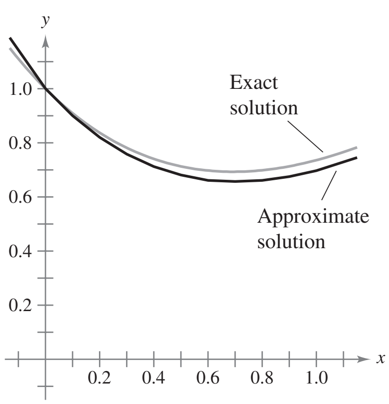


Figure 6.7

Use Euler's Method to approximate the particular solution of the differential equation

$$y' = x - y$$

passing through the point $(0, 1)$. Use a step of $h = 0.1$.

Solution Using $h = 0.1$, $x_0 = 0$, $y_0 = 1$, and $F(x, y) = x - y$, you have $x_0 = 0$, $x_1 = 0.1$, $x_2 = 0.2$, $x_3 = 0.3, \ldots,$ and

$$\begin{aligned}
y_1 &= y_0 + hF(x_0, y_0) = 1 + (0.1)(0 - 1) = 0.9 \\
y_2 &= y_1 + hF(x_1, y_1) = 0.9 + (0.1)(0.1 - 0.9) = 0.82 \\
y_3 &= y_2 + hF(x_2, y_2) = 0.82 + (0.1)(0.2 - 0.82) = 0.758.
\end{aligned}$$

The first ten approximations are shown in the table. You can plot these values to see a graph of the approximate solution, as shown in Figure 6.7.

n	0	1	2	3	4	5	6	7	8	9	10
x_n	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
y_n	1	0.900	0.820	0.758	0.712	0.681	0.663	0.657	0.661	0.675	0.697

Try It   Exploration A

NOTE For the differential equation in Example 6, you can verify the exact solution to be $y = x - 1 + 2e^{-x}$. Figure 6.7 compares this exact solution with the approximate solution obtained in Example 6.

Section 6.2 Differential Equations: Growth and Decay

- **Use separation of variables to solve a simple differential equation.**
- **Use exponential functions to model growth and decay in applied problems.**

Differential Equations

In the preceding section, you learned to analyze visually the solutions of differential equations using slope fields and to approximate solutions numerically using Euler's Method. Analytically, you have learned to solve only two types of differential equations—those of the forms

$$y' = f(x) \quad \text{and} \quad y'' = f(x).$$

In this section, you will learn how to solve a more general type of differential equation. The strategy is to rewrite the equation so that each variable occurs on only one side of the equation. This strategy is called *separation of variables*. (You will study this strategy in detail in Section 6.3.)

NOTE When you integrate both sides of the equation in Example 1, you don't need to add a constant of integration to both sides of the equation. If you did, you would obtain the same result as in Example 1.

$$\int y\,dy = \int 2x\,dx$$
$$\tfrac{1}{2}y^2 + C_2 = x^2 + C_3$$
$$\tfrac{1}{2}y^2 = x^2 + (C_3 - C_2)$$
$$\tfrac{1}{2}y^2 = x^2 + C_1$$

EXAMPLE 1 Solving a Differential Equation

Solve the differential equation $y' = 2x/y$.

Solution

$$y' = \frac{2x}{y} \qquad \text{Write original equation.}$$
$$yy' = 2x \qquad \text{Multiply both sides by } y.$$
$$\int yy'\,dx = \int 2x\,dx \qquad \text{Integrate with respect to } x.$$
$$\int y\,dy = \int 2x\,dx \qquad dy = y'\,dx$$
$$\frac{1}{2}y^2 = x^2 + C_1 \qquad \text{Apply Power Rule.}$$
$$y^2 - 2x^2 = C \qquad \text{Rewrite, letting } C = 2C_1.$$

So, the general solution is given by

$$y^2 - 2x^2 = C.$$

You can use implicit differentiation to check this result.

Try It **Exploration A** **Exploration B** **Exploration C**

EXPLORATION

In Example 1, the general solution of the differential equation is

$$y^2 - 2x^2 = C.$$

Use a graphing utility to sketch several particular solutions—those given by $C = \pm 2$, $C = \pm 1$, and $C = 0$. Describe the solutions graphically. Is the following statement true of each solution?

The slope of the graph at the point (x, y) is equal to twice the ratio of x and y.

Explain your reasoning. Are all curves for which this statement is true represented by the general solution?

In practice, most people prefer to use Leibniz notation and differentials when applying separation of variables. The solution of Example 1 is shown below using this notation.

$$\frac{dy}{dx} = \frac{2x}{y}$$
$$y\,dy = 2x\,dx$$
$$\int y\,dy = \int 2x\,dx$$
$$\frac{1}{2}y^2 = x^2 + C_1$$
$$y^2 - 2x^2 = C$$

Growth and Decay Models

In many applications, the rate of change of a variable y is proportional to the value of y. If y is a function of time t, the proportion can be written as shown.

Rate of change of y is proportional to y.

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

The general solution of this differential equation is given in the following theorem.

> **THEOREM 6.1 Exponential Growth and Decay Model**
>
> If y is a differentiable function of t such that $y > 0$ and $y' = ky$, for some constant k, then
>
> $$y = Ce^{kt}.$$
>
> C is the **initial value** of y, and k is the **proportionality constant. Exponential growth** occurs when $k > 0$, and **exponential decay** occurs when $k < 0$.

NOTE Differentiate the function $y = Ce^{kt}$ with respect to t, and verify that $y' = ky$.

Proof

$$y' = ky \qquad \text{Write original equation.}$$

$$\frac{y'}{y} = k \qquad \text{Separate variables.}$$

$$\int \frac{y'}{y}\,dt = \int k\,dt \qquad \text{Integrate with respect to } t.$$

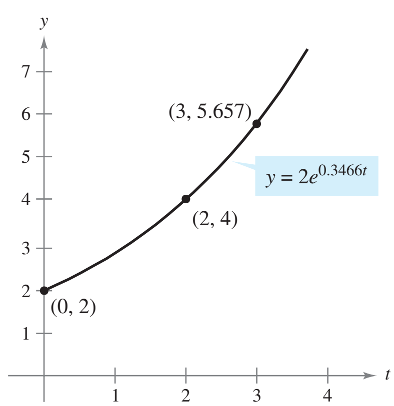
$$\int \frac{1}{y}\,dy = \int k\,dt \qquad dy = y'\,dt$$

$$\ln y = kt + C_1 \qquad \text{Find antiderivative of each side.}$$

$$y = e^{kt}e^{C_1} \qquad \text{Solve for } y.$$

$$y = Ce^{kt} \qquad \text{Let } C = e^{C_1}.$$

So, all solutions of $y' = ky$ are of the form $y = Ce^{kt}$.



If the rate of change of y is proportional to y, then y follows an exponential model.
Figure 6.8

Editable Graph

Select the Animation button below to see that for an exponential decay model, the rate of change of y is proportional to y.

Animation

EXAMPLE 2 Using an Exponential Growth Model

The rate of change of y is proportional to y. When $t = 0$, $y = 2$. When $t = 2$, $y = 4$. What is the value of y when $t = 3$?

Solution Because $y' = ky$, you know that y and t are related by the equation $y = Ce^{kt}$. You can find the values of the constants C and k by applying the initial conditions.

$$2 = Ce^0 \quad \Rightarrow \quad C = 2 \qquad \text{When } t = 0,\ y = 2.$$

$$4 = 2e^{2k} \quad \Rightarrow \quad k = \frac{1}{2}\ln 2 \approx 0.3466 \qquad \text{When } t = 2,\ y = 4.$$

So, the model is $y \approx 2e^{0.3466t}$. When $t = 3$, the value of y is $2e^{0.3466(3)} \approx 5.657$ (see Figure 6.8).

STUDY TIP Using logarithmic properties, note that the value of k in Example 2 can also be written as $\ln(\sqrt{2})$. So, the model becomes $y = 2e^{(\ln\sqrt{2})t}$, which can then be rewritten as $y = 2(\sqrt{2})^t$.

Try It   Exploration A

TECHNOLOGY Most graphing utilities have curve-fitting capabilities that can be used to find models that represent data. Use the *exponential regression* feature of a graphing utility and the information in Example 2 to find a model for the data. How does your model compare with the given model?

Radioactive decay is measured in terms of *half-life*—the number of years required for half of the atoms in a sample of radioactive material to decay. The half-lives of some common radioactive isotopes are shown below.

Uranium (^{238}U)	4,470,000,000 years
Plutonium (^{239}Pu)	24,100 years
Carbon (^{14}C)	5715 years
Radium (^{226}Ra)	1599 years
Einsteinium (^{254}Es)	276 days
Nobelium (^{257}No)	25 seconds

EXAMPLE 3 Radioactive Decay

Suppose that 10 grams of the plutonium isotope Pu-239 was released in the Chernobyl nuclear accident. How long will it take for the 10 grams to decay to 1 gram?

Solution Let y represent the mass (in grams) of the plutonium. Because the rate of decay is proportional to y, you know that

$$y = Ce^{kt}$$

where t is the time in years. To find the values of the constants C and k, apply the initial conditions. Using the fact that $y = 10$ when $t = 0$, you can write

$$10 = Ce^{k(0)} = Ce^0$$

which implies that $C = 10$. Next, using the fact that $y = 5$ when $t = 24{,}100$, you can write

$$\begin{aligned} 5 &= 10e^{k(24{,}100)} \\ \frac{1}{2} &= e^{24{,}100k} \\ \frac{1}{24{,}100}\ln\frac{1}{2} &= k \\ -0.000028761 &\approx k. \end{aligned}$$

So, the model is

$$y = 10e^{-0.000028761t}. \qquad \text{Half-life model}$$

NOTE The exponential decay model in Example 3 could also be written as $y = 10\left(\frac{1}{2}\right)^{t/24{,}100}$. This model is much easier to derive, but for some applications it is not as convenient to use.

To find the time it would take for 10 grams to decay to 1 gram, you can solve for t in the equation

$$1 = 10e^{-0.000028761t}.$$

The solution is approximately 80,059 years.

Try It Exploration A

From Example 3, notice that in an exponential growth or decay problem, it is easy to solve for C when you are given the value of y at $t = 0$. The next example demonstrates a procedure for solving for C and k when you do not know the value of y at $t = 0$.

EXAMPLE 4 Population Growth

Suppose an experimental population of fruit flies increases according to the law of exponential growth. There were 100 flies after the second day of the experiment and 300 flies after the fourth day. Approximately how many flies were in the original population?

Solution Let $y = Ce^{kt}$ be the number of flies at time t, where t is measured in days. Because $y = 100$ when $t = 2$ and $y = 300$ when $t = 4$, you can write

$$100 = Ce^{2k} \quad \text{and} \quad 300 = Ce^{4k}.$$

From the first equation, you know that $C = 100e^{-2k}$. Substituting this value into the second equation produces the following.

$$\begin{aligned} 300 &= 100e^{-2k}e^{4k} \\ 300 &= 100e^{2k} \\ \ln 3 &= 2k \\ \frac{1}{2}\ln 3 &= k \\ 0.5493 &\approx k \end{aligned}$$

So, the exponential growth model is

$$y = Ce^{0.5493t}.$$

To solve for C, reapply the condition $y = 100$ when $t = 2$ and obtain

$$\begin{aligned} 100 &= Ce^{0.5493(2)} \\ C &= 100e^{-1.0986} \approx 33. \end{aligned}$$

So, the original population (when $t = 0$) consisted of approximately $y = C = 33$ flies, as shown in Figure 6.9.

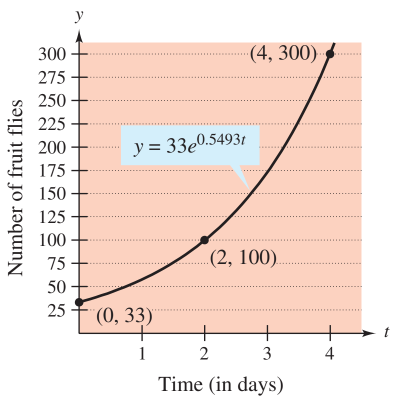


Figure 6.9

Editable Graph

Try It | Exploration A | Open Exploration

EXAMPLE 5 Declining Sales

Four months after it stops advertising, a manufacturing company notices that its sales have dropped from 100,000 units per month to 80,000 units per month. If the sales follow an exponential pattern of decline, what will they be after another 2 months?

Solution Use the exponential decay model $y = Ce^{kt}$, where t is measured in months. From the initial condition $(t = 0)$, you know that $C = 100{,}000$. Moreover, because $y = 80{,}000$ when $t = 4$, you have

$$\begin{aligned} 80{,}000 &= 100{,}000e^{4k} \\ 0.8 &= e^{4k} \\ \ln(0.8) &= 4k \\ -0.0558 &\approx k. \end{aligned}$$

So, after 2 more months $(t = 6)$, you can expect the monthly sales rate to be

$$\begin{aligned} y &\approx 100{,}000e^{-0.0558(6)} \\ &\approx 71{,}500 \text{ units.} \end{aligned}$$

See Figure 6.10.

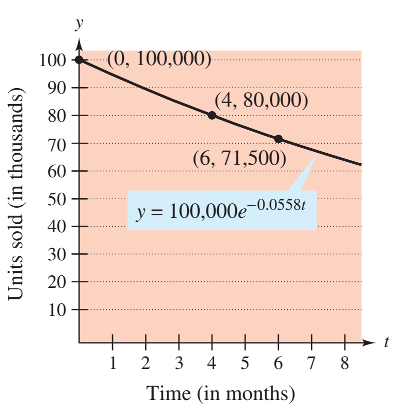


Figure 6.10

Editable Graph

Try It | Exploration A

In Examples 2 through 5, you did not actually have to solve the differential equation

$$y' = ky.$$

(This was done once in the proof of Theorem 6.1.) The next example demonstrates a problem whose solution involves the separation of variables technique. The example concerns **Newton's Law of Cooling,** which states that the rate of change in the temperature of an object is proportional to the difference between the object's temperature and the temperature of the surrounding medium.

EXAMPLE 6 Newton's Law of Cooling

Let y represent the temperature (in °F) of an object in a room whose temperature is kept at a constant 60°. If the object cools from 100° to 90° in 10 minutes, how much longer will it take for its temperature to decrease to 80°?

Solution From Newton's Law of Cooling, you know that the rate of change in y is proportional to the difference between y and 60. This can be written as

$$y' = k(y - 60), \qquad 80 \le y \le 100.$$

To solve this differential equation, use separation of variables, as shown.

$$\frac{dy}{dt} = k(y - 60) \qquad \text{Differential equation}$$

$$\left(\frac{1}{y - 60}\right) dy = k\,dt \qquad \text{Separate variables.}$$

$$\int \frac{1}{y - 60}\,dy = \int k\,dt \qquad \text{Integrate each side.}$$

$$\ln|y - 60| = kt + C_1 \qquad \text{Find antiderivative of each side.}$$

Because $y > 60$, $|y - 60| = y - 60$, and you can omit the absolute value signs. Using exponential notation, you have

$$y - 60 = e^{kt + C_1} \quad \Rightarrow \quad y = 60 + Ce^{kt}. \qquad C = e^{C_1}$$

Using $y = 100$ when $t = 0$, you obtain $100 = 60 + Ce^{k(0)} = 60 + C$, which implies that $C = 40$. Because $y = 90$ when $t = 10$,

$$90 = 60 + 40e^{k(10)}$$
$$30 = 40e^{10k}$$
$$k = \tfrac{1}{10} \ln \tfrac{3}{4} \approx -0.02877.$$

So, the model is

$$y = 60 + 40e^{-0.02877t} \qquad \text{Cooling model}$$

and finally, when $y = 80$, you obtain

$$80 = 60 + 40e^{-0.02877t}$$
$$20 = 40e^{-0.02877t}$$
$$\tfrac{1}{2} = e^{-0.02877t}$$
$$\ln \tfrac{1}{2} = -0.02877t$$
$$t \approx 24.09 \text{ minutes.}$$

So, it will require about 14.09 *more* minutes for the object to cool to a temperature of 80° (see Figure 6.11).

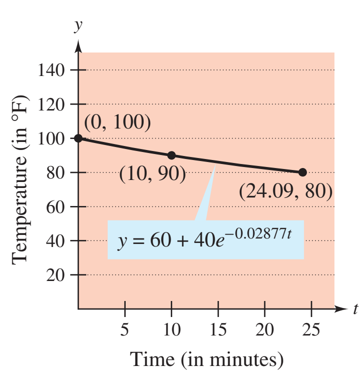


Figure 6.11

Editable Graph

Try It

Exploration A

Section 6.3 Separation of Variables and the Logistic Equation

- **Recognize and solve differential equations that can be solved by separation of variables.**
- **Recognize and solve homogeneous differential equations.**
- **Use differential equations to model and solve applied problems.**
- **Solve and analyze logistic differential equations.**

Separation of Variables

Consider a differential equation that can be written in the form

$$M(x) + N(y)\frac{dy}{dx} = 0$$

where M is a continuous function of x alone and N is a continuous function of y alone. As you saw in the preceding section, for this type of equation, all x terms can be collected with dx and all y terms with dy, and a solution can be obtained by integration. Such equations are said to be **separable,** and the solution procedure is called *separation of variables*. Below are some examples of differential equations that are separable.

Original Differential Equation	*Rewritten with Variables Separated*
$x^2 + 3y\dfrac{dy}{dx} = 0$	$3y\,dy = -x^2\,dx$
$(\sin x)y' = \cos x$	$dy = \cot x\,dx$
$\dfrac{xy'}{e^y + 1} = 2$	$\dfrac{1}{e^y + 1}\,dy = \dfrac{2}{x}\,dx$

EXAMPLE 1 Separation of Variables

Find the general solution of $(x^2 + 4)\dfrac{dy}{dx} = xy$.

Solution To begin, note that $y = 0$ is a solution. To find other solutions, assume that $y \neq 0$ and separate variables as shown.

$$(x^2 + 4)\,dy = xy\,dx \qquad \text{Differential form}$$

$$\frac{dy}{y} = \frac{x}{x^2 + 4}\,dx \qquad \text{Separate variables.}$$

Now, integrate to obtain

$$\int \frac{dy}{y} = \int \frac{x}{x^2 + 4}\,dx \qquad \text{Integrate.}$$

$$\ln|y| = \frac{1}{2}\ln(x^2 + 4) + C_1$$

$$\ln|y| = \ln\sqrt{x^2 + 4} + C_1$$

$$|y| = e^{C_1}\sqrt{x^2 + 4}$$

$$y = \pm e^{C_1}\sqrt{x^2 + 4}.$$

Because $y = 0$ is also a solution, you can write the general solution as

$$y = C\sqrt{x^2 + 4}. \qquad \text{General solution}$$

NOTE Be sure to check your solutions throughout this chapter. In Example 1, you can check the solution $y = C\sqrt{x^2 + 4}$ by differentiating and substituting into the original equation.

$$(x^2 + 4)\frac{dy}{dx} = xy$$

$$(x^2 + 4)\frac{Cx}{\sqrt{x^2 + 4}} \stackrel{?}{=} x\left(C\sqrt{x^2 + 4}\right)$$

$$Cx\sqrt{x^2 + 4} = Cx\sqrt{x^2 + 4}$$

So, the solution checks.

Try It | Exploration A | Open Exploration

The editable graph feature below allows you to edit the graph of a function.

Editable Graph

In some cases it is not feasible to write the general solution in the explicit form $y = f(x)$. The next example illustrates such a solution. Implicit differentiation can be used to verify this solution.

FOR FURTHER INFORMATION For an example (from engineering) of a differential equation that is separable, see the article "Designing a Rose Cutter" by J. S. Hartzler in *The College Mathematics Journal*.

MathArticle

EXAMPLE 2 Finding a Particular Solution

Given the initial condition $y(0) = 1$, find the particular solution of the equation

$$xy\,dx + e^{-x^2}(y^2 - 1)\,dy = 0.$$

Solution Note that $y = 0$ is a solution of the differential equation—but this solution does not satisfy the initial condition. So, you can assume that $y \neq 0$. To separate variables, you must rid the first term of y and the second term of e^{-x^2}. So, you should multiply by e^{x^2}/y and obtain the following.

$$\begin{aligned} xy\,dx + e^{-x^2}(y^2 - 1)\,dy &= 0 \\ e^{-x^2}(y^2 - 1)\,dy &= -xy\,dx \\ \int \left(y - \frac{1}{y}\right) dy &= \int -xe^{x^2}\,dx \\ \frac{y^2}{2} - \ln|y| &= -\frac{1}{2}e^{x^2} + C \end{aligned}$$

From the initial condition $y(0) = 1$, you have $\frac{1}{2} - 0 = -\frac{1}{2} + C$, which implies that $C = 1$. So, the particular solution has the implicit form

$$\frac{y^2}{2} - \ln|y| = -\frac{1}{2}e^{x^2} + 1$$

$$y^2 - \ln y^2 + e^{x^2} = 2.$$

You can check this by differentiating and rewriting to get the original equation.

Try It Exploration A

EXAMPLE 3 Finding a Particular Solution Curve

Find the equation of the curve that passes through the point $(1, 3)$ and has a slope of y/x^2 at any point (x, y).

Solution Because the slope of the curve is given by y/x^2, you have

$$\frac{dy}{dx} = \frac{y}{x^2}$$

with the initial condition $y(1) = 3$. Separating variables and integrating produces

$$\int \frac{dy}{y} = \int \frac{dx}{x^2}, \quad y \neq 0$$

$$\ln|y| = -\frac{1}{x} + C_1$$

$$y = e^{-(1/x) + C_1} = Ce^{-1/x}.$$

Because $y = 3$ when $x = 1$, it follows that $3 = Ce^{-1}$ and $C = 3e$. So, the equation of the specified curve is

$$y = (3e)e^{-1/x} = 3e^{(x-1)/x}, \quad x > 0.$$

See Figure 6.12.

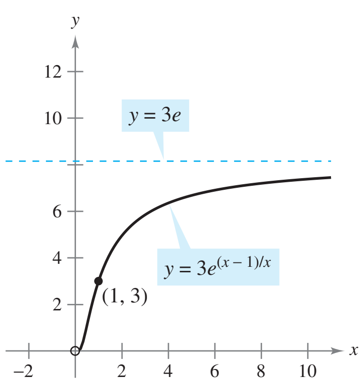


Figure 6.12

Editable Graph

Try It Exploration A

Homogeneous Differential Equations

Some differential equations that are not separable in x and y can be made separable by a change of variables. This is true for differential equations of the form $y' = f(x, y)$, where f is a **homogeneous function.** The function given by $f(x, y)$ is **homogeneous of degree n** if

$$f(tx, ty) = t^n f(x, y) \qquad \text{Homogeneous function of degree } n$$

where n is a real number.

NOTE The notation $f(x, y)$ is used to denote a function of two variables in much the same way as $f(x)$ denotes a function of one variable. You will study functions of two variables in detail in Chapter 13.

EXAMPLE 4 Verifying Homogeneous Functions

a. $f(x, y) = x^2y - 4x^3 + 3xy^2$ is a homogeneous function of degree 3 because

$$\begin{aligned} f(tx, ty) &= (tx)^2(ty) - 4(tx)^3 + 3(tx)(ty)^2 \\ &= t^3(x^2y) - t^3(4x^3) + t^3(3xy^2) \\ &= t^3(x^2y - 4x^3 + 3xy^2) \\ &= t^3 f(x, y). \end{aligned}$$

b. $f(x, y) = xe^{x/y} + y\sin(y/x)$ is a homogeneous function of degree 1 because

$$\begin{aligned} f(tx, ty) &= txe^{tx/ty} + ty \sin\frac{ty}{tx} \\ &= t\left(xe^{x/y} + y\sin\frac{y}{x}\right) \\ &= tf(x, y). \end{aligned}$$

c. $f(x, y) = x + y^2$ is *not* a homogeneous function because

$$f(tx, ty) = tx + t^2y^2 = t(x + ty^2) \neq t^n(x + y^2).$$

d. $f(x, y) = x/y$ is a homogeneous function of degree 0 because

$$f(tx, ty) = \frac{tx}{ty} = t^0\frac{x}{y}.$$

Try It Exploration A

Definition of Homogeneous Differential Equation

A **homogeneous differential equation** is an equation of the form

$$M(x, y)\,dx + N(x, y)\,dy = 0$$

where M and N are homogeneous functions of the same degree.

EXAMPLE 5 Testing for Homogeneous Differential Equations

a. $(x^2 + xy)\,dx + y^2\,dy = 0$ is homogeneous of degree 2.

b. $x^3\,dx = y^3\,dy$ is homogeneous of degree 3.

c. $(x^2 + 1)\,dx + y^2\,dy = 0$ is *not* a homogeneous differential equation.

Try It Exploration A

To solve a homogeneous differential equation by the method of separation of variables, use the following change of variables theorem.

THEOREM 6.2 Change of Variables for Homogeneous Equations

If $M(x, y)\,dx + N(x, y)\,dy = 0$ is homogeneous, then it can be transformed into a differential equation whose variables are separable by the substitution

$$y = vx$$

where v is a differentiable function of x.

EXAMPLE 6 Solving a Homogeneous Differential Equation

Find the general solution of

$$(x^2 - y^2)\,dx + 3xy\,dy = 0.$$

STUDY TIP The substitution $y = vx$ will yield a differential equation that is separable with respect to the variables x and v. You must write your final solution, however, in terms of x and y.

Solution Because $(x^2 - y^2)$ and $3xy$ are both homogeneous of degree 2, let $y = vx$ to obtain $dy = x\,dv + v\,dx$. Then, by substitution, you have

$$\begin{aligned}
(x^2 - v^2x^2)\,dx + 3x(vx)\overbrace{(x\,dv + v\,dx)}^{dy} &= 0\\
(x^2 + 2v^2x^2)\,dx + 3x^3v\,dv &= 0\\
x^2(1 + 2v^2)\,dx + x^2(3vx)\,dv &= 0.
\end{aligned}$$

Dividing by x^2 and separating variables produces

$$\begin{aligned}
(1 + 2v^2)\,dx &= -3vx\,dv\\
\int \frac{dx}{x} &= \int \frac{-3v}{1 + 2v^2}\,dv\\
\ln|x| &= -\frac{3}{4}\ln(1 + 2v^2) + C_1\\
4\ln|x| &= -3\ln(1 + 2v^2) + \ln|C|\\
\ln x^4 &= \ln|C(1 + 2v^2)^{-3}|\\
x^4 &= C(1 + 2v^2)^{-3}.
\end{aligned}$$

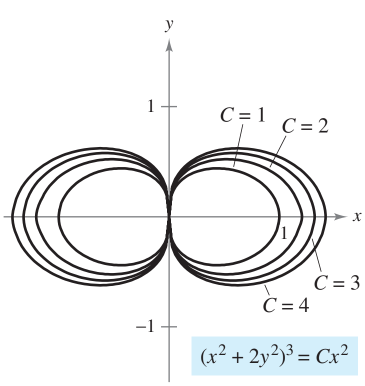
Substituting for v produces the following general solution.

$$\begin{aligned}
x^4 &= C\left[1 + 2\left(\frac{y}{x}\right)^2\right]^{-3}\\
\left(1 + \frac{2y^2}{x^2}\right)^3 x^4 &= C\\
(x^2 + 2y^2)^3 &= Cx^2 && \text{General solution}
\end{aligned}$$

You can check this by differentiating and rewriting to get the original equation.

Try It

Exploration A



General solutions of
$(x^2 - y^2)\,dx + 3xy\,dy = 0$
Figure 6.13

TECHNOLOGY If you have access to a graphing utility, try using it to graph several of the solutions in Example 6. For instance, Figure 6.13 shows the graphs of

$$(x^2 + 2y^2)^3 = Cx^2$$

for $C = 1, 2, 3$, and 4.

Applications

EXAMPLE 7 Wildlife Population

The rate of change of the number of coyotes $N(t)$ in a population is directly proportional to $650 - N(t)$, where t is the time in years. When $t = 0$, the population is 300, and when $t = 2$, the population has increased to 500. Find the population when $t = 3$.

Solution Because the rate of change of the population is proportional to $650 - N(t)$, you can write the following differential equation.

$$\frac{dN}{dt} = k(650 - N)$$

You can solve this equation using separation of variables.

$$dN = k(650 - N)\,dt \qquad \text{Differential form}$$

$$\frac{dN}{650 - N} = k\,dt \qquad \text{Separate variables.}$$

$$-\ln|650 - N| = kt + C_1 \qquad \text{Integrate.}$$

$$\ln|650 - N| = -kt - C_1$$

$$650 - N = e^{-kt - C_1} \qquad \text{Assume } N < 650.$$

$$N = 650 - Ce^{-kt} \qquad \text{General solution}$$

Using $N = 300$ when $t = 0$, you can conclude that $C = 350$, which produces

$$N = 650 - 350e^{-kt}.$$

Then, using $N = 500$ when $t = 2$, it follows that

$$500 = 650 - 350e^{-2k} \quad \Longrightarrow \quad e^{-2k} = \frac{3}{7} \quad \Longrightarrow \quad k \approx 0.4236.$$

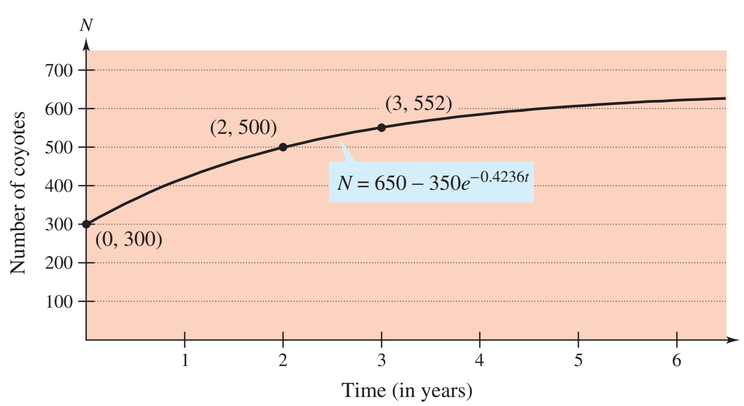
So, the model for the coyote population is

$$N = 650 - 350e^{-0.4236t}. \qquad \text{Model for population}$$

When $t = 3$, you can approximate the population to be

$$N = 650 - 350e^{-0.4236(3)} \approx 552 \text{ coyotes.}$$

The model for the population is shown in Figure 6.14.

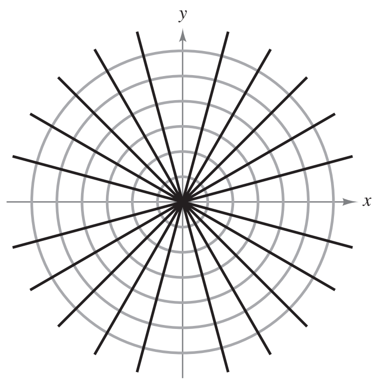


Editable Graph

Figure 6.14

Try It

Exploration A



Each line $y = Kx$ is an orthogonal trajectory to the family of circles.
Figure 6.15

A common problem in electrostatics, thermodynamics, and hydrodynamics involves finding a family of curves, each of which is orthogonal to all members of a given family of curves. For example, Figure 6.15 shows a family of circles

$$x^2 + y^2 = C \qquad \text{Family of circles}$$

each of which intersects the lines in the family

$$y = Kx \qquad \text{Family of lines}$$

at right angles. Two such families of curves are said to be **mutually orthogonal,** and each curve in one of the families is called an **orthogonal trajectory** of the other family. In electrostatics, lines of force are orthogonal to the *equipotential curves*. In thermodynamics, the flow of heat across a plane surface is orthogonal to the *isothermal curves*. In hydrodynamics, the flow (stream) lines are orthogonal trajectories of the *velocity potential curves*.

EXAMPLE 8 Finding Orthogonal Trajectories

Describe the orthogonal trajectories for the family of curves given by

$$y = \frac{C}{x}$$

for $C \neq 0$. Sketch several members of each family.

Solution First, solve the given equation for C and write $xy = C$. Then, by differentiating implicitly with respect to x, you obtain the differential equation

$$xy' + y = 0 \qquad \text{Differential equation}$$

$$x\frac{dy}{dx} = -y$$

$$\frac{dy}{dx} = -\frac{y}{x}. \qquad \text{Slope of given family}$$

Because y' represents the slope of the given family of curves at (x, y), it follows that the orthogonal family has the negative reciprocal slope x/y. So,

$$\frac{dy}{dx} = \frac{x}{y}. \qquad \text{Slope of orthogonal family}$$

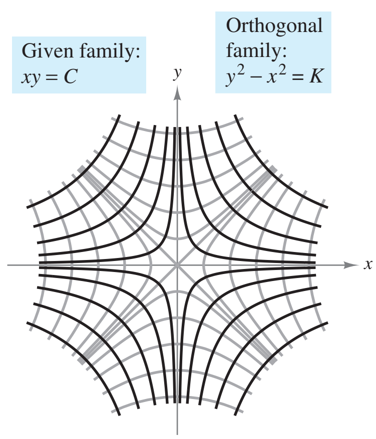
Now you can find the orthogonal family by separating variables and integrating.

$$\int y\,dy = \int x\,dx$$

$$\frac{y^2}{2} = \frac{x^2}{2} + C_1$$

$$y^2 - x^2 = K$$

The centers are at the origin, and the transverse axes are vertical for $K > 0$ and horizontal for $K < 0$. If $k = 0$, the orthogonal trajectories are the lines $y = \pm x$. If $K \neq 0$, the orthogonal trajectories are hyperbolas. Several trajectories are shown in Figure 6.16.



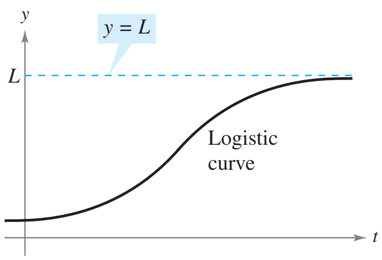
Orthogonal trajectories
Figure 6.16

Try It Exploration A

Logistic Differential Equation

In Section 6.2, the exponential growth model is derived from the fact that the rate of change of a variable y is proportional to the value of y. You observed that the differential equation $dy/dt = ky$ has the general solution $y = Ce^{kt}$. Exponential growth is unlimited, but when describing a population, there often exists some upper limit L past which growth cannot occur. This upper limit L is called the **carrying capacity,** which is the maximum population $y(t)$ that can be sustained or supported as time t increases. A model that is often used for this type of growth is the **logistic differential equation**

$$\frac{dy}{dt} = ky\left(1 - \frac{y}{L}\right)$$ Logistic differential equation



Note that as $t \to \infty$, $y \to L$.
Figure 6.17

where k and L are positive constants. A population that satisfies this equation does not grow without bound, but approaches the carrying capacity L as t increases.

From the equation, you can see that if y is between 0 and the carrying capacity L, then $dy/dt > 0$, and the population increases. If k is greater than L, then $dy/dt < 0$, and the population decreases. The graph of the function y is called the *logistic curve*, as shown in Figure 6.17.

EXAMPLE 9 Deriving the General Solution

Solve the logistic differential equation $\dfrac{dy}{dt} = ky\left(1 - \dfrac{y}{L}\right)$.

Solution Begin by separating variables.

$$\frac{dy}{dt} = ky\left(1 - \frac{y}{L}\right)$$ Write differential equation.

$$\frac{1}{y(1 - y/L)}dy = kdt$$ Separate variables.

$$\int \frac{1}{y(1 - y/L)}dy = \int kdt$$ Integrate each side.

$$\int \left(\frac{1}{y} + \frac{1}{L - y}\right)dy = \int kdt$$ Rewrite left side using partial fractions.

$$\ln|y| - \ln|L - y| = kt + C$$ Find antiderivative of each side.

$$\ln\left|\frac{L - y}{y}\right| = -kt - C$$ Multiply each side by -1 and simplify.

$$\left|\frac{L - y}{y}\right| = e^{-kt-C} = e^{-C}e^{-kt}$$ Exponentiate each side.

$$\frac{L - y}{y} = be^{-kt}$$ Let $\pm e^{-C} = b$.

Solving this equation for y produces $y = \dfrac{L}{1 + be^{-kt}}$.

Try It | Exploration A | Exploration B

EXPLORATION

Use a graphing utility to investigate the effects of the values of L, b, and k on the graph of

$$y = \frac{L}{1 + be^{-kt}}.$$

Include some examples to support your results.

From Example 9, you can conclude that all solutions of the logistic differential equation are of the general form

$$y = \frac{L}{1 + be^{-kt}}.$$

EXAMPLE 10 Solving a Logistic Differential Equation

A state game commission releases 40 elk into a game refuge. After 5 years, the elk population is 104. The commission believes that the environment can support no more than 4000 elk. The growth rate of the elk population p is

$$\frac{dp}{dt} = kp\left(1 - \frac{p}{4000}\right), \quad 40 \le p \le 4000$$

where t is the number of years.

a. Write a model for the elk population in terms of t.

b. Graph the slope field of the differential equation and the solution that passes through the point $(0, 40)$.

c. Use the model to estimate the elk population after 15 years.

d. Find the limit of the model as $t \to \infty$.

Solution

a. You know that $L = 4000$. So, the solution of the equation is of the form

$$p = \frac{4000}{1 + be^{-kt}}.$$

Because $p(0) = 40$, you can solve for b as shown.

$$40 = \frac{4000}{1 + be^{-k(0)}}$$

$$40 = \frac{4000}{1 + b} \quad \Rightarrow \quad b = 99$$

Then, because $p = 104$ when $t = 5$, you can solve for k.

$$104 = \frac{4000}{1 + 99e^{-k(5)}} \quad \Rightarrow \quad k \approx 0.194$$

So, a model for the elk population is given by $p = \dfrac{4000}{1 + 99e^{-0.194t}}$.

b. Using a graphing utility, you can graph the slope field of

$$\frac{dp}{dt} = 0.194p\left(1 - \frac{p}{4000}\right)$$

and the solution that passes through $(0, 40)$, as shown in Figure 6.18.

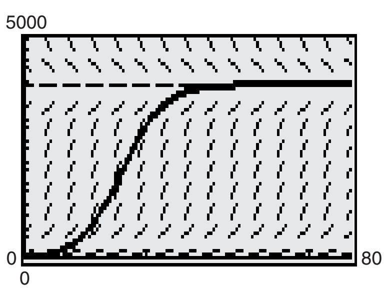


Figure 6.18

c. To estimate the elk population after 15 years, substitute 15 for t in the model

$$p = \frac{4000}{1 + 99e^{-0.194(15)}} \qquad \text{Substitute 15 for } t.$$

$$= \frac{4000}{1 + 99e^{-2.91}} \approx 626 \qquad \text{Simplify.}$$

d. As t increases without bound, the denominator of $\dfrac{4000}{1 + 99e^{-0.194t}}$ gets closer to 1.

So, $\displaystyle\lim_{t\to\infty} \frac{4000}{1 + 99e^{-0.194t}} = 4000.$

Try It **Exploration A**

EXPLORATION

Explain what happens if $p(0) = L$.

Section 6.4 First-Order Linear Differential Equations

- **Solve a first-order linear differential equation.**
- **Solve a Bernoulli differential equation.**
- **Use linear differential equations to solve applied problems.**

First-Order Linear Differential Equations

In this section, you will see how to solve a very important class of first-order differential equations—first-order linear differential equations.

Definition of First-Order Linear Differential Equation

A first-order linear differential equation is an equation of the form

$$\frac{dy}{dx} + P(x)y = Q(x)$$

where P and Q are continuous functions of x. This first-order linear differential equation is said to be in **standard form.**

NOTE It is instructive to see why the integrating factor helps solve a linear differential equation of the form $y' + P(x)y = Q(x)$. When both sides of the equation are multiplied by the integrating factor $u(x) = e^{\int P(x)\,dx}$, the left-hand side becomes the derivative of a product.

$$y'e^{\int P(x)\,dx} + P(x)ye^{\int P(x)\,dx} = Q(x)e^{\int P(x)\,dx}$$
$$\left[ye^{\int P(x)\,dx}\right]' = Q(x)e^{\int P(x)\,dx}$$

Integrating both sides of this second equation and dividing by $u(x)$ produces the general solution.

To solve a linear differential equation, write it in standard form to identify the functions $P(x)$ and $Q(x)$. Then integrate $P(x)$ and form the expression

$$u(x) = e^{\int P(x)\,dx} \qquad \text{Integrating factor}$$

which is called an **integrating factor.** The general solution of the equation is

$$y = \frac{1}{u(x)}\int Q(x)u(x)\,dx. \qquad \text{General solution}$$

EXAMPLE 1 Solving a Linear Differential Equation

Find the general solution of

$$y' + y = e^x.$$

Solution

For this equation, $P(x) = 1$ and $Q(x) = e^x$. So, the integrating factor is

$$\begin{aligned} u(x) &= e^{\int P(x)\,dx} && \text{Integrating factor} \\ &= e^{\int dx} \\ &= e^x. \end{aligned}$$

This implies that the general solution is

$$\begin{aligned} y &= \frac{1}{u(x)}\int Q(x)u(x)\,dx \\ &= \frac{1}{e^x}\int e^x(e^x)\,dx \\ &= e^{-x}\left(\frac{1}{2}e^{2x} + C\right) \\ &= \frac{1}{2}e^x + Ce^{-x}. && \text{General solution} \end{aligned}$$

Try It Exploration A

ANNA JOHNSON PELL WHEELER (1883–1966)

Anna Johnson Pell Wheeler was awarded a master's degree from the University of Iowa for her thesis *The Extension of Galois Theory to Linear Differential Equations* in 1904. Influenced by David Hilbert, she worked on integral equations while studying infinite linear spaces.

THEOREM 6.3 Solution of a First-Order Linear Differential Equation

An integrating factor for the first-order linear differential equation

$$y' + P(x)y = Q(x)$$

is $u(x) = e^{\int P(x)\,dx}$. The solution of the differential equation is

$$ye^{\int P(x)\,dx} = \int Q(x)e^{\int P(x)\,dx}\,dx + C.$$

STUDY TIP Rather than memorizing the formula in Theorem 6.3, just remember that multiplication by the integrating factor $e^{\int P(x)\,dx}$ converts the left side of the differential equation into the derivative of the product $ye^{\int P(x)\,dx}$.

EXAMPLE 2 Solving a First-Order Linear Differential Equation

Find the general solution of

$$xy' - 2y = x^2.$$

Solution The standard form of the given equation is

$$y' + P(x)y = Q(x)$$

$$y' - \left(\frac{2}{x}\right)y = x. \qquad \text{Standard form}$$

So, $P(x) = -2/x$, and you have

$$\int P(x)\,dx = -\int \frac{2}{x}\,dx$$

$$= -\ln x^2$$

$$e^{\int P(x)\,dx} = e^{-\ln x^2}$$

$$= \frac{1}{e^{\ln x^2}}$$

$$= \frac{1}{x^2}. \qquad \text{Integrating factor}$$

So, multiplying each side of the standard form by $1/x^2$ yields

$$\frac{y'}{x^2} - \frac{2y}{x^3} = \frac{1}{x}$$

$$\frac{d}{dx}\left[\frac{y}{x^2}\right] = \frac{1}{x}$$

$$\frac{y}{x^2} = \int \frac{1}{x}\,dx$$

$$\frac{y}{x^2} = \ln|x| + C$$

$$y = x^2(\ln|x| + C). \qquad \text{General solution}$$

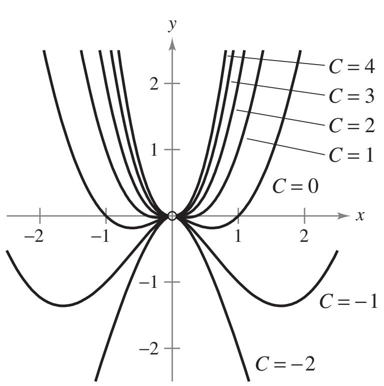


Figure 6.19

Several solution curves (for $C = -2, -1, 0, 1, 2, 3,$ and 4) are shown in Figure 6.19.

Editable Graph

Try It

Exploration A

EXAMPLE 3 Solving a First-Order Linear Differential Equation

Find the general solution of

$$y' - y \tan t = 1, \quad -\frac{\pi}{2} < t < \frac{\pi}{2}.$$

Solution The equation is already in the standard form $y' + P(t)y = Q(t)$. So, $P(t) = -\tan t$, and

$$\int P(t)\,dt = -\int \tan t\,dt = \ln|\cos t|$$

which implies that the integrating factor is

$$e^{\int P(t)\,dt} = e^{\ln|\cos t|}$$
$$= |\cos t|. \qquad \text{Integrating factor}$$

A quick check shows that $\cos t$ is also an integrating factor. So, multiplying $y' - y \tan t = 1$ by $\cos t$ produces

$$\frac{d}{dt}[y \cos t] = \cos t$$
$$y \cos t = \int \cos t\,dt$$
$$y \cos t = \sin t + C$$
$$y = \tan t + C \sec t. \qquad \text{General solution}$$

Several solution curves are shown in Figure 6.20.

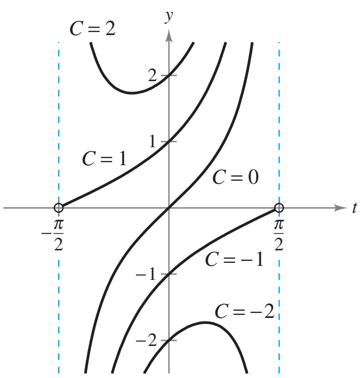


Figure 6.20

Editable Graph

Try It | Exploration A | Open Exploration

Bernoulli Equation

A well-known nonlinear equation that reduces to a linear one with an appropriate substitution is the **Bernoulli equation,** named after James Bernoulli (1654–1705).

$$y' + P(x)y = Q(x)y^n \qquad \text{Bernoulli equation}$$

This equation is linear if $n = 0$, and has separable variables if $n = 1$. So, in the following development, assume that $n \neq 0$ and $n \neq 1$. Begin by multiplying by y^{-n} and $(1 - n)$ to obtain

$$y^{-n}y' + P(x)y^{1-n} = Q(x)$$
$$(1-n)y^{-n}y' + (1-n)P(x)y^{1-n} = (1-n)Q(x)$$
$$\frac{d}{dx}[y^{1-n}] + (1-n)P(x)y^{1-n} = (1-n)Q(x)$$

which is a linear equation in the variable y^{1-n}. Letting $z = y^{1-n}$ produces the linear equation

$$\frac{dz}{dx} + (1-n)P(x)z = (1-n)Q(x).$$

Finally, by Theorem 6.3, the *general solution of the Bernoulli equation* is

$$y^{1-n}e^{\int(1-n)P(x)\,dx} = \int (1-n)Q(x)e^{\int(1-n)P(x)\,dx}\,dx + C.$$

EXAMPLE 4 Solving a Bernoulli Equation

Find the general solution of

$$y' + xy = xe^{-x^2}y^{-3}.$$

Solution For this Bernoulli equation, let $n = -3$, and use the substitution

$$z = y^4 \qquad \text{Let } z = y^{1-n} = y^{1-(-3)}.$$
$$z' = 4y^3y'. \qquad \text{Differentiate.}$$

Multiplying the original equation by $4y^3$ produces

$$y' + xy = xe^{-x^2}y^{-3} \qquad \text{Write original equation.}$$
$$4y^3y' + 4xy^4 = 4xe^{-x^2} \qquad \text{Multiply each side by } 4y^3.$$
$$z' + 4xz = 4xe^{-x^2}. \qquad \text{Linear equation: } z' + P(x)z = Q(x)$$

This equation is linear in z. Using $P(x) = 4x$ produces

$$\int P(x)\,dx = \int 4x\,dx$$
$$= 2x^2$$

which implies that e^{2x^2} is an integrating factor. Multiplying the linear equation by this factor produces

$$z' + 4xz = 4xe^{-x^2} \qquad \text{Linear equation}$$
$$z'e^{2x^2} + 4xze^{2x^2} = 4xe^{x^2} \qquad \text{Multiply by integrating factor.}$$
$$\frac{d}{dx}\left[ze^{2x^2}\right] = 4xe^{x^2} \qquad \text{Write left side as derivative.}$$
$$ze^{2x^2} = \int 4xe^{x^2}\,dx \qquad \text{Integrate each side.}$$
$$ze^{2x^2} = 2e^{x^2} + C$$
$$z = 2e^{-x^2} + Ce^{-2x^2}. \qquad \text{Divide each side by } e^{2x^2}.$$

Finally, substituting $z = y^4$, the general solution is

$$y^4 = 2e^{-x^2} + Ce^{-2x^2}. \qquad \text{General solution}$$

Try It **Exploration A**

So far you have studied several types of first-order differential equations. Of these, the separable variables case is usually the simplest, and solution by an integrating factor is ordinarily used only as a last resort.

Summary of First-Order Differential Equations

Method	*Form of Equation*
1. Separable variables:	$M(x)\,dx + N(y)\,dy = 0$
2. Homogeneous:	$M(x, y)\,dx + N(x, y)\,dy = 0$, where M and N are nth-degree homogeneous
3. Linear:	$y' + P(x)y = Q(x)$
4. Bernoulli equation:	$y' + P(x)y = Q(x)y^n$

Applications

One type of problem that can be described in terms of a differential equation involves chemical mixtures, as illustrated in the next example.

EXAMPLE 5 A Mixture Problem

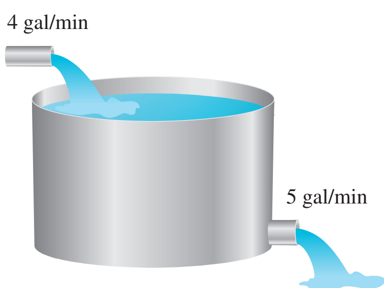


Figure 6.21

A tank contains 50 gallons of a solution composed of 90% water and 10% alcohol. A second solution containing 50% water and 50% alcohol is added to the tank at the rate of 4 gallons per minute. As the second solution is being added, the tank is being drained at a rate of 5 gallons per minute, as shown in Figure 6.21. Assuming the solution in the tank is stirred constantly, how much alcohol is in the tank after 10 minutes?

Solution Let y be the number of gallons of alcohol in the tank at any time t. You know that $y = 5$ when $t = 0$. Because the number of gallons of solution in the tank at any time is $50 - t$, and the tank loses 5 gallons of solution per minute, it must lose

$$\left(\frac{5}{50 - t}\right)y$$

gallons of alcohol per minute. Furthermore, because the tank is gaining 2 gallons of alcohol per minute, the rate of change of alcohol in the tank is given by

$$\frac{dy}{dt} = 2 - \left(\frac{5}{50 - t}\right)y \quad \Longrightarrow \quad \frac{dy}{dt} + \left(\frac{5}{50 - t}\right)y = 2.$$

To solve this linear equation, let $P(t) = 5/(50 - t)$ and obtain

$$\int P(t)\,dt = \int \frac{5}{50 - t}\,dt = -5 \ln|50 - t|.$$

Because $t < 50$, you can drop the absolute value signs and conclude that

$$e^{\int P(t)\,dt} = e^{-5\ln(50 - t)} = \frac{1}{(50 - t)^5}.$$

So, the general solution is

$$\frac{y}{(50 - t)^5} = \int \frac{2}{(50 - t)^5}\,dt = \frac{1}{2(50 - t)^4} + C$$

$$y = \frac{50 - t}{2} + C(50 - t)^5.$$

Because $y = 5$ when $t = 0$, you have

$$5 = \frac{50}{2} + C(50)^5 \quad \Longrightarrow \quad -\frac{20}{50^5} = C$$

which means that the particular solution is

$$y = \frac{50 - t}{2} - 20\left(\frac{50 - t}{50}\right)^5.$$

Finally, when $t = 10$, the amount of alcohol in the tank is

$$y = \frac{50 - 10}{2} - 20\left(\frac{50 - 10}{50}\right)^5 \approx 13.45 \text{ gal}$$

which represents a solution containing 33.6% alcohol.

Try It Exploration A

In most falling-body problems discussed so far in the text, air resistance has been neglected. The next example includes this factor. In the example, the air resistance on the falling object is assumed to be proportional to its velocity v. If g is the gravitational constant, the downward force F on a falling object of mass m is given by the difference $mg - kv$. But by Newton's Second Law of Motion, you know that

$$\begin{aligned} F &= ma \\ &= m(dv/dt) \end{aligned}$$

which yields the following differential equation.

$$m\frac{dv}{dt} = mg - kv \quad \Rightarrow \quad \frac{dv}{dt} + \frac{k}{m}v = g$$

EXAMPLE 6 A Falling Object with Air Resistance

An object of mass m is dropped from a hovering helicopter. Find its velocity as a function of time t, assuming that the air resistance is proportional to the velocity of the object.

Solution The velocity v satisfies the equation

$$\frac{dv}{dt} + \frac{kv}{m} = g$$

where g is the gravitational constant and k is the constant of proportionality. Letting $b = k/m$, you can *separate variables* to obtain

$$\begin{aligned} dv &= (g - bv)\, dt \\ \int \frac{dv}{g - bv} &= \int dt \\ -\frac{1}{b}\ln|g - bv| &= t + C_1 \\ \ln|g - bv| &= -bt - bC_1 \\ g - bv &= Ce^{-bt}. \end{aligned}$$

Because the object was dropped, $v = 0$ when $t = 0$; so $g = C$, and it follows that

$$-bv = -g + ge^{-bt} \quad \Rightarrow \quad v = \frac{g - ge^{-bt}}{b} = \frac{mg}{k}(1 - e^{-kt/m}).$$

Try It **Exploration A**

NOTE Notice in Example 6 that the velocity approaches a limit of mg/k as a result of the air resistance. For falling-body problems in which air resistance is neglected, the velocity increases without bound.

A simple electric circuit consists of electric current I (in amperes), a resistance R (in ohms), an inductance L (in henrys), and a constant electromotive force E (in volts), as shown in Figure 6.22. According to Kirchhoff's Second Law, if the switch S is closed when $t = 0$, the applied electromotive force (voltage) is equal to the sum of the voltage drops in the rest of the circuit. This in turn means that the current I satisfies the differential equation

$$L\frac{dI}{dt} + RI = E.$$

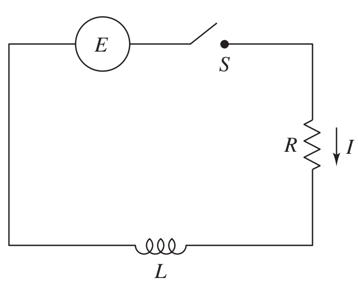


Figure 6.22

EXAMPLE 7 An Electric Circuit Problem

Find the current I as a function of time t (in seconds), given that I satisfies the differential equation $L(dI/dt) + RI = \sin 2t$, where R and L are nonzero constants.

TECHNOLOGY The integral in Example 7 was found using symbolic algebra software. If you have access to *Derive*, *Maple*, *Mathcad*, *Mathematica*, or the *TI-89*, try using it to integrate

$$\frac{1}{L}\int e^{(R/L)t} \sin 2t\, dt.$$

In Chapter 8 you will learn how to integrate functions of this type using integration by parts.

Solution In standard form, the given linear equation is

$$\frac{dI}{dt} + \frac{R}{L} I = \frac{1}{L} \sin 2t.$$

Let $P(t) = R/L$, so that $e^{\int P(t)\,dt} = e^{(R/L)t}$, and, by Theorem 6.3,

$$\begin{aligned} Ie^{(R/L)t} &= \frac{1}{L}\int e^{(R/L)t} \sin 2t\, dt \\ &= \frac{1}{4L^2 + R^2} e^{(R/L)t}(R \sin 2t - 2L \cos 2t) + C. \end{aligned}$$

So the general solution is

$$I = e^{-(R/L)t}\left[\frac{1}{4L^2 + R^2} e^{(R/L)t}(R \sin 2t - 2L \cos 2t) + C\right]$$

$$I = \frac{1}{4L^2 + R^2}(R \sin 2t - 2L \cos 2t) + Ce^{-(R/L)t}.$$

Try It Exploration A

Section 7.1 Area of a Region Between Two Curves

- **Find the area of a region between two curves using integration.**
- **Find the area of a region between intersecting curves using integration.**
- **Describe integration as an accumulation process.**

Area of a Region Between Two Curves

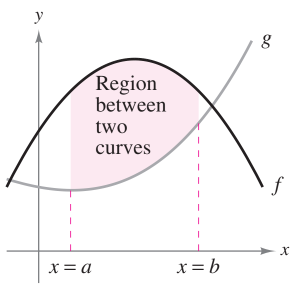
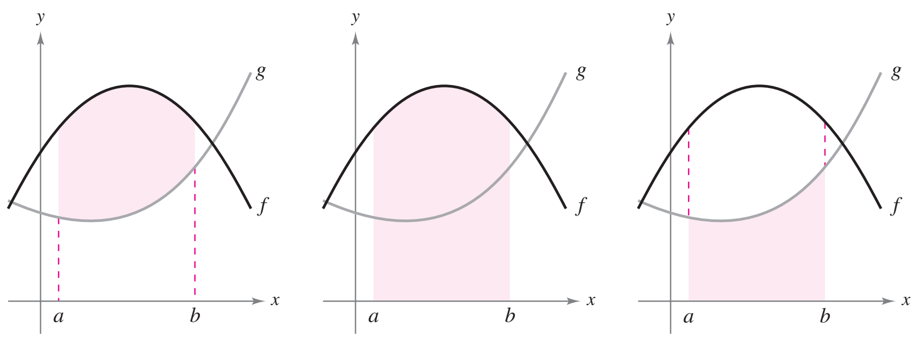


Figure 7.1

With a few modifications you can extend the application of definite integrals from the area of a region *under* a curve to the area of a region *between* two curves. Consider two functions f and g that are continuous on the interval $[a, b]$. If, as in Figure 7.1, the graphs of both f and g lie above the x-axis, and the graph of g lies below the graph of f, you can geometrically interpret the area of the region between the graphs as the area of the region under the graph of g subtracted from the area of the region under the graph of f, as shown in Figure 7.2.



Area of region between f and g = Area of region under f − Area of region under g

$$\int_a^b [f(x) - g(x)]\,dx = \int_a^b f(x)\,dx - \int_a^b g(x)\,dx$$

Figure 7.2

Animation

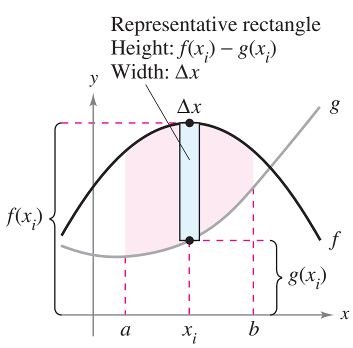


Figure 7.3

To verify the reasonableness of the result shown in Figure 7.2, you can partition the interval $[a, b]$ into n subintervals, each of width Δx. Then, as shown in Figure 7.3, sketch a **representative rectangle** of width Δx and height $f(x_i) - g(x_i)$, where x_i is in the ith interval. The area of this representative rectangle is

$$\Delta A_i = (\text{height})(\text{width}) = [f(x_i) - g(x_i)]\Delta x.$$

By adding the areas of the n rectangles and taking the limit as $\|\Delta\| \to 0$ $(n \to \infty)$, you obtain

$$\lim_{n\to\infty} \sum_{i=1}^{n} [f(x_i) - g(x_i)]\Delta x.$$

Because f and g are continuous on $[a, b]$, $f - g$ is also continuous on $[a, b]$ and the limit exists. So, the area of the given region is

$$\begin{aligned}\text{Area} &= \lim_{n\to\infty} \sum_{i=1}^{n} [f(x_i) - g(x_i)]\Delta x \\ &= \int_a^b [f(x) - g(x)]\,dx.\end{aligned}$$

Area of a Region Between Two Curves

If f and g are continuous on $[a, b]$ and $g(x) \le f(x)$ for all x in $[a, b]$, then the area of the region bounded by the graphs of f and g and the vertical lines $x = a$ and $x = b$ is

$$A = \int_a^b [f(x) - g(x)]\,dx.$$

In Figure 7.1, the graphs of f and g are shown above the x-axis. This, however, is not necessary. The same integrand $[f(x) - g(x)]$ can be used as long as f and g are continuous and $g(x) \le f(x)$ for all x in the interval $[a, b]$. This result is summarized graphically in Figure 7.4.

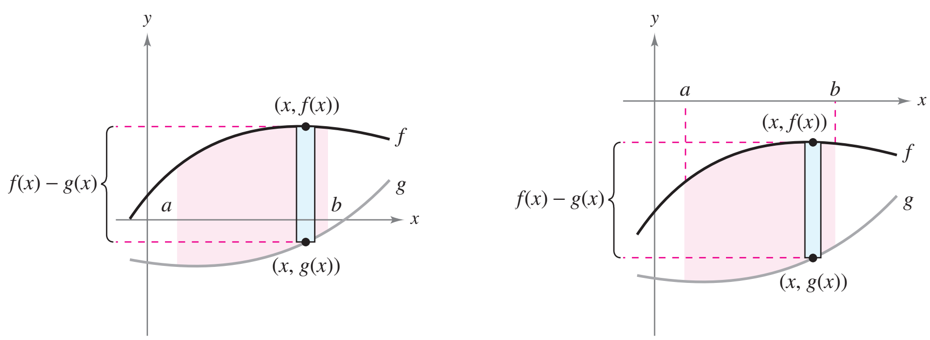


Figure 7.4

NOTE The height of a representative rectangle is $f(x) - g(x)$ regardless of the relative position of the x-axis, as shown in Figure 7.4.

Representative rectangles are used throughout this chapter in various applications of integration. A vertical rectangle (of width Δx) implies integration with respect to x, whereas a horizontal rectangle (of width Δy) implies integration with respect to y.

EXAMPLE 1 Finding the Area of a Region Between Two Curves

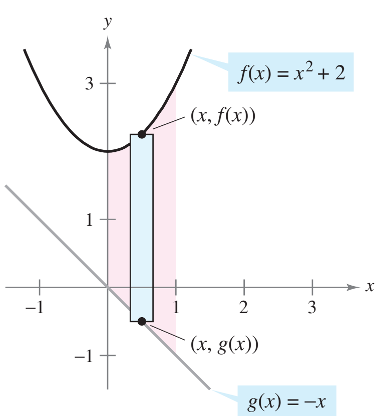
Find the area of the region bounded by the graphs of $y = x^2 + 2$, $y = -x$, $x = 0$, and $x = 1$.

Solution Let $g(x) = -x$ and $f(x) = x^2 + 2$. Then $g(x) \le f(x)$ for all x in $[0, 1]$, as shown in Figure 7.5. So, the area of the representative rectangle is

$$\begin{aligned}\Delta A &= [f(x) - g(x)]\,\Delta x \\ &= [(x^2 + 2) - (-x)]\,\Delta x\end{aligned}$$

and the area of the region is

$$\begin{aligned}A = \int_a^b [f(x) - g(x)]\,dx &= \int_0^1 [(x^2 + 2) - (-x)]\,dx \\ &= \left[\frac{x^3}{3} + \frac{x^2}{2} + 2x\right]_0^1 \\ &= \frac{1}{3} + \frac{1}{2} + 2 \\ &= \frac{17}{6}.\end{aligned}$$



Region bounded by the graph of f, the graph of g, $x = 0$, and $x = 1$
Figure 7.5

Editable Graph

Try It

Exploration A

Exploration B

Area of a Region Between Intersecting Curves

In Example 1, the graphs of $f(x) = x^2 + 2$ and $g(x) = -x$ do not intersect, and the values of a and b are given explicitly. A more common problem involves the area of a region bounded by two *intersecting* graphs, where the values of a and b must be calculated.

EXAMPLE 2 A Region Lying Between Two Intersecting Graphs

Find the area of the region bounded by the graphs of $f(x) = 2 - x^2$ and $g(x) = x$.

Solution In Figure 7.6, notice that the graphs of f and g have two points of intersection. To find the x-coordinates of these points, set $f(x)$ and $g(x)$ equal to each other and solve for x.

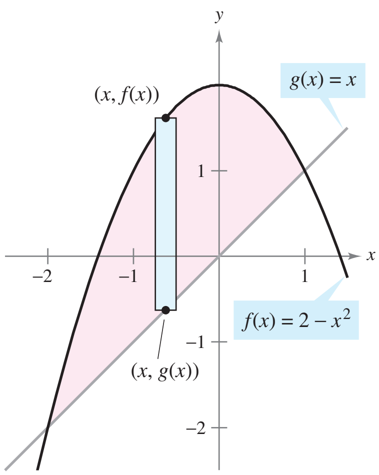
$$2 - x^2 = x \qquad \text{Set } f(x) \text{ equal to } g(x).$$
$$-x^2 - x + 2 = 0 \qquad \text{Write in general form.}$$
$$-(x + 2)(x - 1) = 0 \qquad \text{Factor.}$$
$$x = -2 \text{ or } 1 \qquad \text{Solve for } x.$$

So, $a = -2$ and $b = 1$. Because $g(x) \le f(x)$ for all x in the interval $[-2, 1]$, the representative rectangle has an area of

$$\begin{aligned}\Delta A &= [f(x) - g(x)]\,\Delta x\\ &= [(2 - x^2) - x]\,\Delta x\end{aligned}$$

and the area of the region is

$$A = \int_{-2}^{1} [(2 - x^2) - x]\,dx = \left[-\frac{x^3}{3} - \frac{x^2}{2} + 2x\right]_{-2}^{1} = \frac{9}{2}.$$



Region bounded by the graph of f and the graph of g
Figure 7.6

Editable Graph

Try It Exploration A

EXAMPLE 3 A Region Lying Between Two Intersecting Graphs

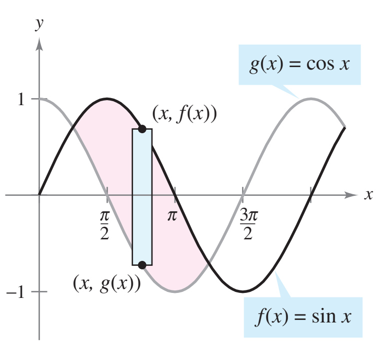
The sine and cosine curves intersect infinitely many times, bounding regions of equal areas, as shown in Figure 7.7. Find the area of one of these regions.

Solution

$$\sin x = \cos x \qquad \text{Set } f(x) \text{ equal to } g(x).$$
$$\frac{\sin x}{\cos x} = 1 \qquad \text{Divide each side by } \cos x.$$
$$\tan x = 1 \qquad \text{Trigonometric identity}$$
$$x = \frac{\pi}{4} \text{ or } \frac{5\pi}{4}, \quad 0 \le x \le 2\pi \qquad \text{Solve for } x.$$

So, $a = \pi/4$ and $b = 5\pi/4$. Because $\sin x \ge \cos x$ for all x in the interval $[\pi/4, 5\pi/4]$, the area of the region is

$$A = \int_{\pi/4}^{5\pi/4} [\sin x - \cos x]\,dx = \left[-\cos x - \sin x\right]_{\pi/4}^{5\pi/4} = 2\sqrt{2}.$$



One of the regions bounded by the graphs of the sine and cosine functions
Figure 7.7

Editable Graph

Try It Exploration A

If two curves intersect at more than two points, then to find the area of the region between the curves, you must find all points of intersection and check to see which curve is above the other in each interval determined by these points.

EXAMPLE 4 Curves That Intersect at More Than Two Points

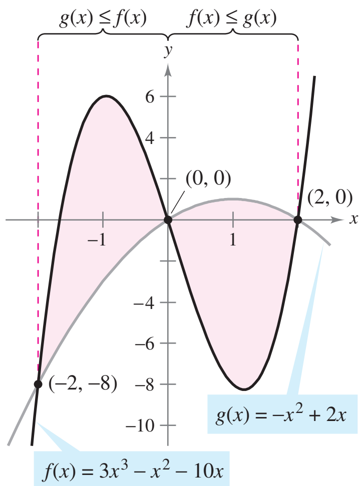
Find the area of the region between the graphs of $f(x) = 3x^3 - x^2 - 10x$ and $g(x) = -x^2 + 2x$.

Solution Begin by setting $f(x)$ and $g(x)$ equal to each other and solving for x. This yields the x-values at each point of intersection of the two graphs.

$$3x^3 - x^2 - 10x = -x^2 + 2x \qquad \text{Set } f(x) \text{ equal to } g(x).$$
$$3x^3 - 12x = 0 \qquad \text{Write in general form.}$$
$$3x(x-2)(x+2) = 0 \qquad \text{Factor.}$$
$$x = -2, 0, 2 \qquad \text{Solve for } x.$$

So, the two graphs intersect when $x = -2, 0,$ and 2. In Figure 7.8, notice that $g(x) \le f(x)$ on the interval $[-2, 0]$. However, the two graphs switch at the origin, and $f(x) \le g(x)$ on the interval $[0, 2]$. So, you need two integrals—one for the interval $[-2, 0]$ and one for the interval $[0, 2]$.

$$\begin{aligned} A &= \int_{-2}^{0} [f(x) - g(x)]\,dx + \int_{0}^{2} [g(x) - f(x)]\,dx \\ &= \int_{-2}^{0} (3x^3 - 12x)\,dx + \int_{0}^{2} (-3x^3 + 12x)\,dx \\ &= \left[\frac{3x^4}{4} - 6x^2\right]_{-2}^{0} + \left[\frac{-3x^4}{4} + 6x^2\right]_{0}^{2} \\ &= -(12 - 24) + (-12 + 24) = 24 \end{aligned}$$



On $[-2, 0]$, $g(x) \le f(x)$, and on $[0, 2]$, $f(x) \le g(x)$
Figure 7.8

Editable Graph

Try It Exploration A Open Exploration

NOTE In Example 4, notice that you obtain an incorrect result if you integrate from -2 to 2. Such integration produces

$$\int_{-2}^{2} [f(x) - g(x)]\,dx = \int_{-2}^{2} (3x^3 - 12x)\,dx = 0.$$

If the graph of a function of y is a boundary of a region, it is often convenient to use representative rectangles that are *horizontal* and find the area by integrating with respect to y. In general, to determine the area between two curves, you can use

$$A = \int_{x_1}^{x_2} \underbrace{[(\text{top curve}) - (\text{bottom curve})]}_{\text{in variable } x}\,dx \qquad \text{Vertical rectangles}$$

$$A = \int_{y_1}^{y_2} \underbrace{[(\text{right curve}) - (\text{left curve})]}_{\text{in variable } y}\,dy \qquad \text{Horizontal rectangles}$$

where (x_1, y_1) and (x_2, y_2) are either adjacent points of intersection of the two curves involved or points on the specified boundary lines.

Technology

EXAMPLE 5 Horizontal Representative Rectangles

Find the area of the region bounded by the graphs of $x = 3 - y^2$ and $x = y + 1$.

Solution Consider

$$g(y) = 3 - y^2 \quad \text{and} \quad f(y) = y + 1.$$

These two curves intersect when $y = -2$ and $y = 1$, as shown in Figure 7.9. Because $f(y) \le g(y)$ on this interval, you have

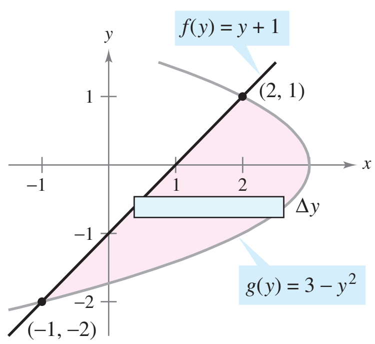
$$\Delta A = [g(y) - f(y)]\,\Delta y = [(3 - y^2) - (y + 1)]\,\Delta y.$$

So, the area is

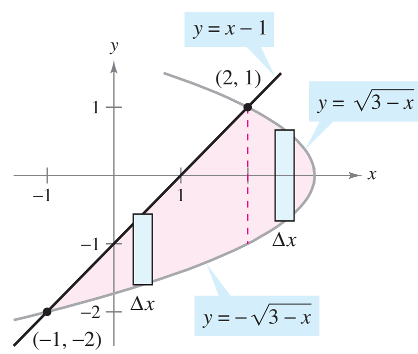
$$\begin{aligned} A &= \int_{-2}^{1} [(3 - y^2) - (y + 1)]\,dy \\ &= \int_{-2}^{1} (-y^2 - y + 2)\,dy \\ &= \left[\frac{-y^3}{3} - \frac{y^2}{2} + 2y\right]_{-2}^{1} \\ &= \left(-\frac{1}{3} - \frac{1}{2} + 2\right) - \left(\frac{8}{3} - 2 - 4\right) \\ &= \frac{9}{2}. \end{aligned}$$

Try It

Exploration A



Horizontal rectangles (integration with respect to y)
Figure 7.9



Vertical rectangles (integration with respect to x)
Figure 7.10

Editable Graph

In Example 5, notice that by integrating with respect to y you need only one integral. If you had integrated with respect to x, you would have needed two integrals because the upper boundary would have changed at $x = 2$, as shown in Figure 7.10.

$$\begin{aligned} A &= \int_{-1}^{2} \left[(x - 1) + \sqrt{3 - x}\right] dx + \int_{2}^{3} \left(\sqrt{3 - x} + \sqrt{3 - x}\right) dx \\ &= \int_{-1}^{2} [x - 1 + (3 - x)^{1/2}]\,dx + 2\int_{2}^{3} (3 - x)^{1/2}\,dx \\ &= \left[\frac{x^2}{2} - x - \frac{(3 - x)^{3/2}}{3/2}\right]_{-1}^{2} - 2\left[\frac{(3 - x)^{3/2}}{3/2}\right]_{2}^{3} \\ &= \left(2 - 2 - \frac{2}{3}\right) - \left(\frac{1}{2} + 1 - \frac{16}{3}\right) - 2(0) + 2\left(\frac{2}{3}\right) \\ &= \frac{9}{2} \end{aligned}$$

Integration as an Accumulation Process

In this section, the integration formula for the area between two curves was developed by using a rectangle as the *representative element*. For each new application in the remaining sections of this chapter, an appropriate representative element will be constructed using precalculus formulas you already know. Each integration formula will then be obtained by summing or accumulating these representative elements.

Known precalculus formula ⇨ Representative element ⇨ New integration formula

For example, in this section the area formula was developed as follows.

$$A = (\text{height})(\text{width}) \quad \Rightarrow \quad \Delta A = [f(x) - g(x)]\,\Delta x \quad \Rightarrow \quad A = \int_a^b [f(x) - g(x)]\,dx$$

EXAMPLE 6 Describing Integration as an Accumulation Process

Find the area of the region bounded by the graph of $y = 4 - x^2$ and the x-axis. Describe the integration as an accumulation process.

Solution The area of the region is given by

$$A = \int_{-2}^{2} (4 - x^2)\,dx.$$

You can think of the integration as an accumulation of the areas of the rectangles formed as the representative rectangle slides from $x = -2$ to $x = 2$, as shown in Figure 7.11.

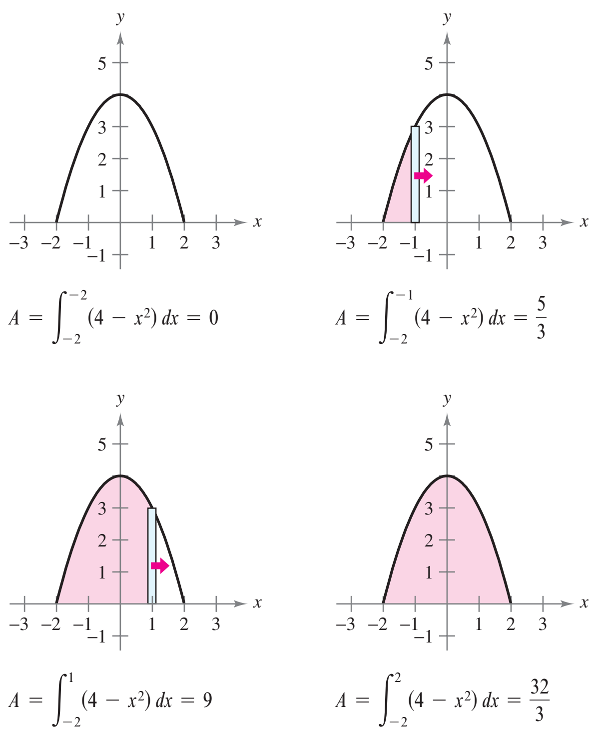


Figure 7.11

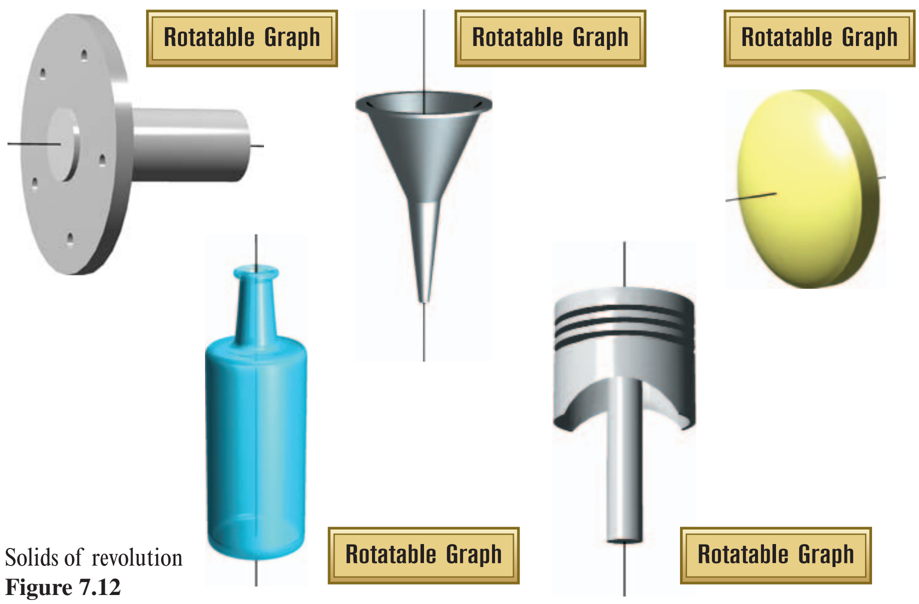
Try It **Exploration A**

Section 7.2 Volume: The Disk Method

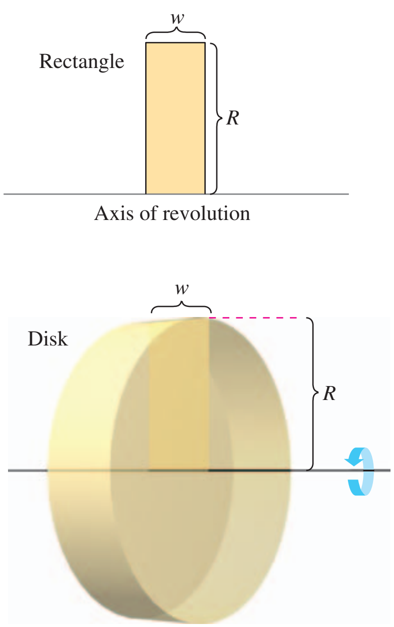
- **Find the volume of a solid of revolution using the disk method.**
- **Find the volume of a solid of revolution using the washer method.**
- **Find the volume of a solid with known cross sections.**

The Disk Method

In Chapter 4 we mentioned that area is only one of the *many* applications of the definite integral. Another important application is its use in finding the volume of a three-dimensional solid. In this section you will study a particular type of three-dimensional solid—one whose cross sections are similar. Solids of revolution are used commonly in engineering and manufacturing. Some examples are axles, funnels, pills, bottles, and pistons, as shown in Figure 7.12.



Solids of revolution
Figure 7.12



Volume of a disk: $\pi R^2 w$
Figure 7.13

Rotatable Graph

If a region in the plane is revolved about a line, the resulting solid is a **solid of revolution,** and the line is called the **axis of revolution.** The simplest such solid is a right circular cylinder or **disk,** which is formed by revolving a rectangle about an axis adjacent to one side of the rectangle, as shown in Figure 7.13. The volume of such a disk is

$$\begin{aligned}\text{Volume of disk} &= (\text{area of disk})(\text{width of disk})\\ &= \pi R^2 w\end{aligned}$$

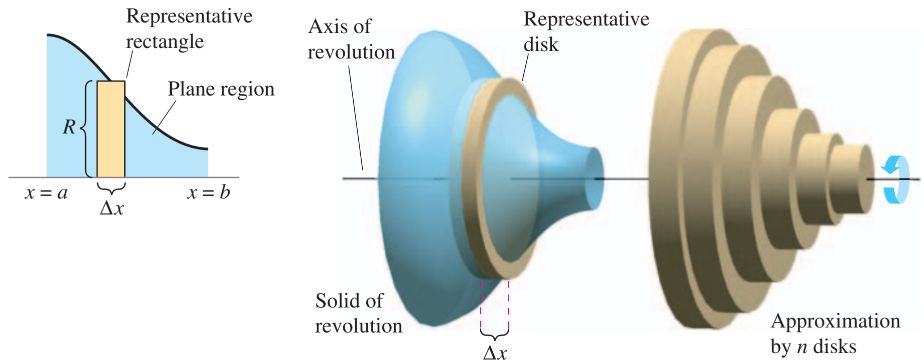
where R is the radius of the disk and w is the width.

To see how to use the volume of a disk to find the volume of a general solid of revolution, consider a solid of revolution formed by revolving the plane region in Figure 7.14 about the indicated axis. To determine the volume of this solid, consider a representative rectangle in the plane region. When this rectangle is revolved about the axis of revolution, it generates a representative disk whose volume is

$$\Delta V = \pi R^2 \Delta x.$$

Approximating the volume of the solid by n such disks of width Δx and radius $R(x_i)$ produces

$$\begin{aligned}\text{Volume of solid} &\approx \sum_{i=1}^{n} \pi[R(x_i)]^2\,\Delta x\\ &= \pi\sum_{i=1}^{n} [R(x_i)]^2\,\Delta x.\end{aligned}$$



Disk method
Figure 7.14

Rotatable Graph

This approximation appears to become better and better as $\|\Delta\| \to 0$ $(n \to \infty)$. So, you can define the volume of the solid as

$$\text{Volume of solid} = \lim_{\|\Delta\| \to 0} \pi \sum_{i=1}^{n} [R(x_i)]^2 \, \Delta x = \pi \int_a^b [R(x)]^2 \, dx.$$

Schematically, the disk method looks like this.

Known Precalculus Formula		*Representative Element*		*New Integration Formula*
Volume of disk $V = \pi R^2 w$	⇒	$\Delta V = \pi [R(x_i)]^2 \, \Delta x$	⇒	Solid of revolution $V = \pi \int_a^b [R(x)]^2 \, dx$

A similar formula can be derived if the axis of revolution is vertical.

The Disk Method

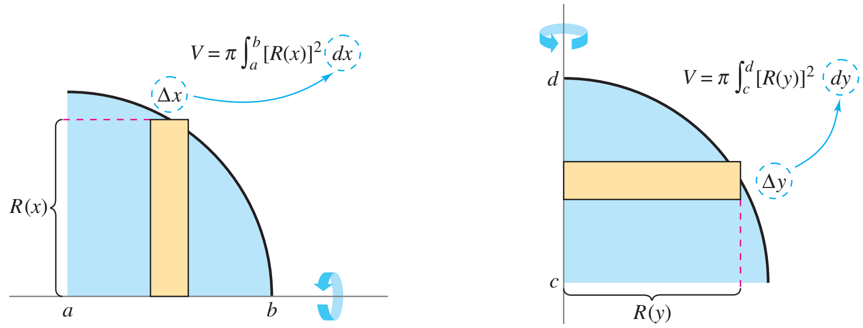
To find the volume of a solid of revolution with the **disk method,** use one of the following, as shown in Figure 7.15.

Horizontal Axis of Revolution

$$\text{Volume} = V = \pi \int_a^b [R(x)]^2 \, dx$$

Vertical Axis of Revolution

$$\text{Volume} = V = \pi \int_c^d [R(y)]^2 \, dy$$



Horizontal axis of revolution   Vertical axis of revolution
Figure 7.15

NOTE In Figure 7.15, note that you can determine the variable of integration by placing a representative rectangle in the plane region "perpendicular" to the axis of revolution. If the width of the rectangle is Δx, integrate with respect to x, and if the width of the rectangle is Δy, integrate with respect to y.

The simplest application of the disk method involves a plane region bounded by the graph of f and the x-axis. If the axis of revolution is the x-axis, the radius $R(x)$ is simply $f(x)$.

EXAMPLE 1 Using the Disk Method

Find the volume of the solid formed by revolving the region bounded by the graph of

$$f(x) = \sqrt{\sin x}$$

and the x-axis $(0 \le x \le \pi)$ about the x-axis.

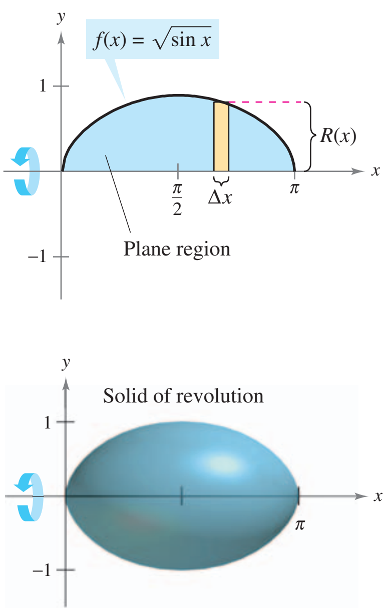


Figure 7.16

Rotatable Graph

Solution From the representative rectangle in the upper graph in Figure 7.16, you can see that the radius of this solid is

$$\begin{aligned} R(x) &= f(x) \\ &= \sqrt{\sin x}. \end{aligned}$$

So, the volume of the solid of revolution is

$$\begin{aligned} V = \pi \int_a^b [R(x)]^2\,dx &= \pi \int_0^{\pi} \left(\sqrt{\sin x}\right)^2 dx && \text{Apply disk method.} \\ &= \pi \int_0^{\pi} \sin x\,dx && \text{Simplify.} \\ &= \pi \Big[-\cos x\Big]_0^{\pi} && \text{Integrate.} \\ &= \pi(1 + 1) \\ &= 2\pi. \end{aligned}$$

Try It Exploration A Exploration B

EXAMPLE 2 Revolving About a Line That Is Not a Coordinate Axis

Find the volume of the solid formed by revolving the region bounded by

$$f(x) = 2 - x^2$$

and $g(x) = 1$ about the line $y = 1$, as shown in Figure 7.17.

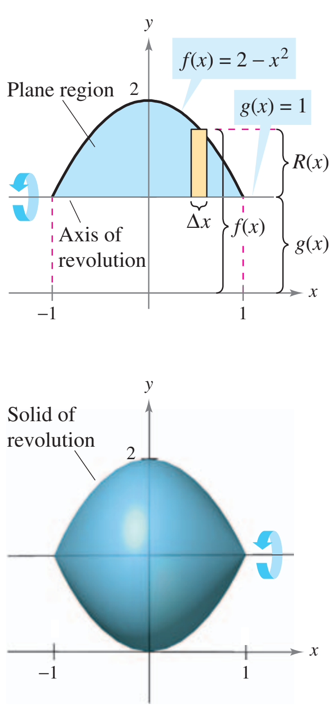


Figure 7.17

Rotatable Graph

Solution By equating $f(x)$ and $g(x)$, you can determine that the two graphs intersect when $x = \pm 1$. To find the radius, subtract $g(x)$ from $f(x)$.

$$\begin{aligned} R(x) &= f(x) - g(x) \\ &= (2 - x^2) - 1 \\ &= 1 - x^2 \end{aligned}$$

Finally, integrate between -1 and 1 to find the volume.

$$\begin{aligned} V = \pi \int_a^b [R(x)]^2\,dx &= \pi \int_{-1}^{1} (1 - x^2)^2\,dx && \text{Apply disk method.} \\ &= \pi \int_{-1}^{1} (1 - 2x^2 + x^4)\,dx && \text{Simplify.} \\ &= \pi \left[x - \frac{2x^3}{3} + \frac{x^5}{5}\right]_{-1}^{1} && \text{Integrate.} \\ &= \frac{16\pi}{15} \end{aligned}$$

Try It Exploration A

The Washer Method

The disk method can be extended to cover solids of revolution with holes by replacing the representative disk with a representative **washer.** The washer is formed by revolving a rectangle about an axis, as shown in Figure 7.18. If r and R are the inner and outer radii of the washer and w is the width of the washer, the volume is given by

$$\text{Volume of washer} = \pi(R^2 - r^2)w.$$

To see how this concept can be used to find the volume of a solid of revolution, consider a region bounded by an **outer radius** $R(x)$ and an **inner radius** $r(x)$, as shown in Figure 7.19. If the region is revolved about its axis of revolution, the volume of the resulting solid is given by

$$V = \pi\int_a^b ([R(x)]^2 - [r(x)]^2)\,dx. \qquad \text{Washer method}$$

Note that the integral involving the inner radius represents the volume of the hole and is *subtracted* from the integral involving the outer radius.

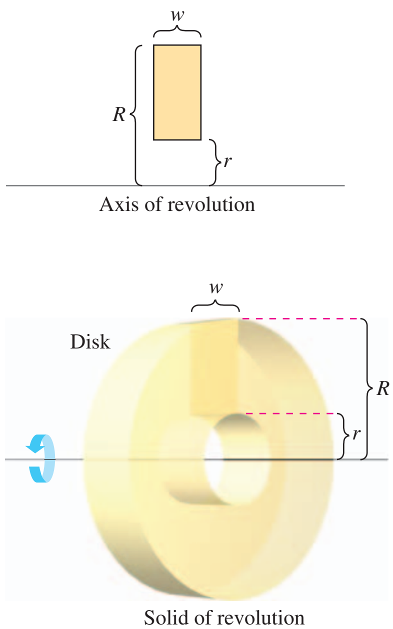


Figure 7.18

Rotatable Graph

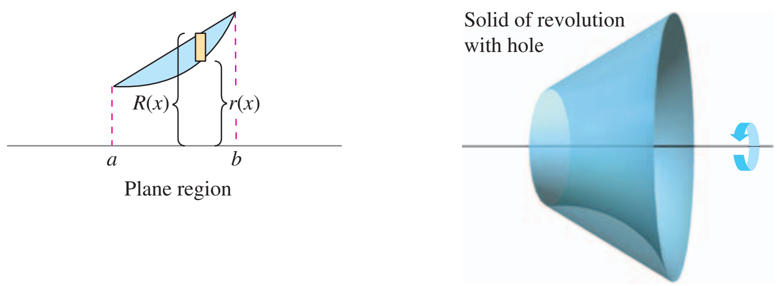


Figure 7.19

Rotatable Graph

EXAMPLE 3 Using the Washer Method

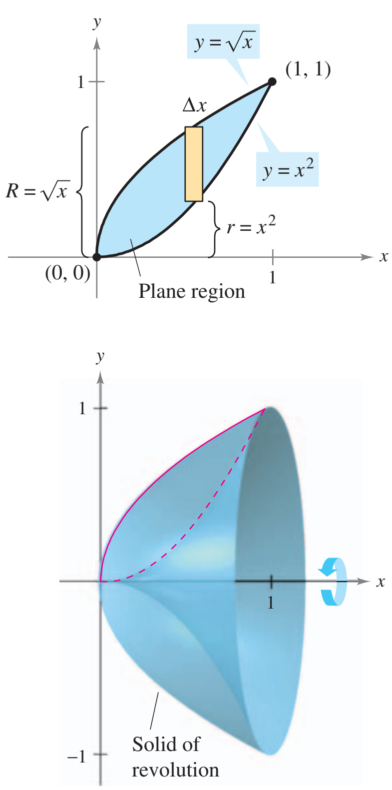
Find the volume of the solid formed by revolving the region bounded by the graphs of $y = \sqrt{x}$ and $y = x^2$ about the x-axis, as shown in Figure 7.20.

Solution In Figure 7.20, you can see that the outer and inner radii are as follows.

$$R(x) = \sqrt{x} \qquad \text{Outer radius}$$
$$r(x) = x^2 \qquad \text{Inner radius}$$

Integrating between 0 and 1 produces

$$\begin{aligned} V &= \pi\int_a^b ([R(x)]^2 - [r(x)]^2)\,dx && \text{Apply washer method.}\\ &= \pi\int_0^1 \left[\left(\sqrt{x}\right)^2 - (x^2)^2\right] dx \\ &= \pi\int_0^1 (x - x^4)\,dx && \text{Simplify.}\\ &= \pi\left[\frac{x^2}{2} - \frac{x^5}{5}\right]_0^1 && \text{Integrate.}\\ &= \frac{3\pi}{10}. \end{aligned}$$



Solid of revolution
Figure 7.20

Rotatable Graph

Try It Exploration A

In each example so far, the axis of revolution has been *horizontal* and you have integrated with respect to x. In the next example, the axis of revolution is *vertical* and you integrate with respect to y. In this example, you need two separate integrals to compute the volume.

EXAMPLE 4 Integrating with Respect to y, Two-Integral Case

Find the volume of the solid formed by revolving the region bounded by the graphs of $y = x^2 + 1$, $y = 0$, $x = 0$, and $x = 1$ about the y-axis, as shown in Figure 7.21.

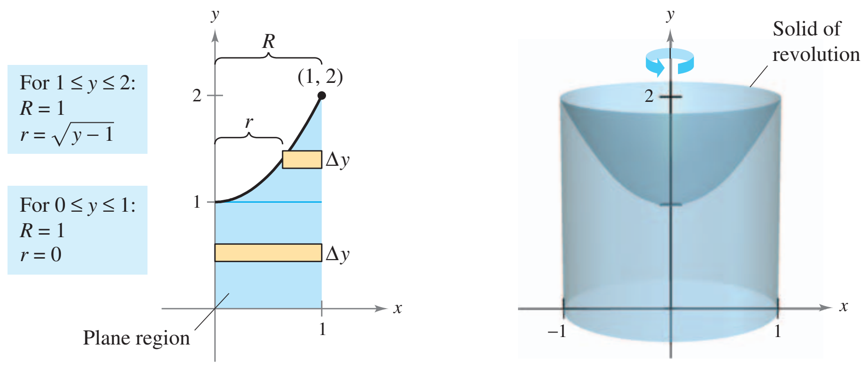


Figure 7.21

Rotatable Graph

Solution For the region shown in Figure 7.21, the outer radius is simply $R = 1$. There is, however, no convenient formula that represents the inner radius. When $0 \le y \le 1$, $r = 0$, but when $1 \le y \le 2$, r is determined by the equation $y = x^2 + 1$, which implies that $r = \sqrt{y-1}$.

$$r(y) = \begin{cases} 0, & 0 \le y \le 1 \\ \sqrt{y-1}, & 1 \le y \le 2 \end{cases}$$

Using this definition of the inner radius, you can use two integrals to find the volume.

$$\begin{aligned} V &= \pi \int_0^1 (1^2 - 0^2)\,dy + \pi \int_1^2 \left[1^2 - \left(\sqrt{y-1}\right)^2\right] dy && \text{Apply washer method.} \\ &= \pi \int_0^1 1\,dy + \pi \int_1^2 (2-y)\,dy && \text{Simplify.} \\ &= \pi \Big[y\Big]_0^1 + \pi \left[2y - \frac{y^2}{2}\right]_1^2 && \text{Integrate.} \\ &= \pi + \pi\left(4 - 2 - 2 + \frac{1}{2}\right) = \frac{3\pi}{2} \end{aligned}$$

Note that the first integral $\pi \int_0^1 1\,dy$ represents the volume of a right circular cylinder of radius 1 and height 1. This portion of the volume could have been determined without using calculus.

Try It | Exploration A | Exploration B

TECHNOLOGY Some graphing utilities have the capability to generate (or have built-in software capable of generating) a solid of revolution. If you have access to such a utility, use it to graph some of the solids of revolution described in this section. For instance, the solid in Example 4 might appear like that shown in Figure 7.22.

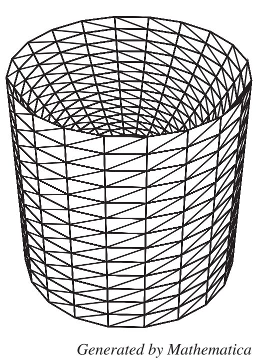
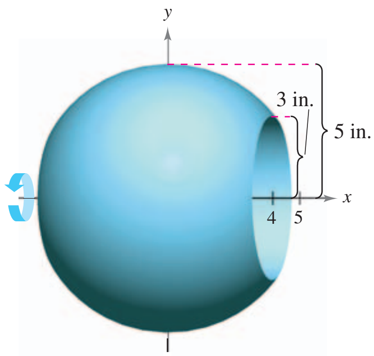
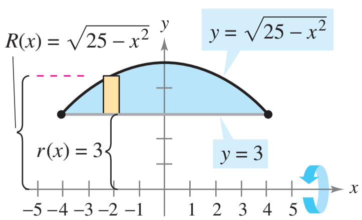


Figure 7.22



Solid of revolution

(a)



Plane region

(b)
Figure 7.23

Rotatable Graph

EXAMPLE 5 Manufacturing

A manufacturer drills a hole through the center of a metal sphere of radius 5 inches, as shown in Figure 7.23(a). The hole has a radius of 3 inches. What is the volume of the resulting metal ring?

Solution You can imagine the ring to be generated by a segment of the circle whose equation is $x^2 + y^2 = 25$, as shown in Figure 7.23(b). Because the radius of the hole is 3 inches, you can let $y = 3$ and solve the equation $x^2 + y^2 = 25$ to determine that the limits of integration are $x = \pm 4$. So, the inner and outer radii are $r(x) = 3$ and $R(x) = \sqrt{25 - x^2}$ and the volume is given by

$$\begin{aligned} V = \pi \int_a^b ([R(x)]^2 - [r(x)]^2)\,dx &= \pi \int_{-4}^{4} \left[\left(\sqrt{25 - x^2}\right)^2 - (3)^2\right] dx \\ &= \pi \int_{-4}^{4} (16 - x^2)\,dx \\ &= \pi \left[16x - \frac{x^3}{3}\right]_{-4}^{4} \\ &= \frac{256\pi}{3} \text{ cubic inches.} \end{aligned}$$

Try It Exploration A Open Exploration

Solids with Known Cross Sections

With the disk method, you can find the volume of a solid having a circular cross section whose area is $A = \pi R^2$. This method can be generalized to solids of any shape, as long as you know a formula for the area of an arbitrary cross section. Some common cross sections are squares, rectangles, triangles, semicircles, and trapezoids.

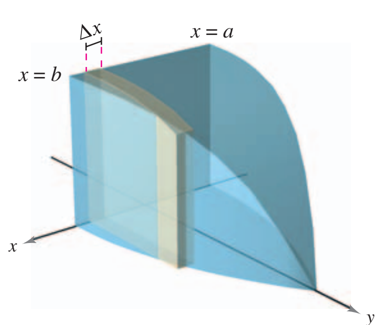
Volumes of Solids with Known Cross Sections

1. For cross sections of area $A(x)$ taken perpendicular to the x-axis,

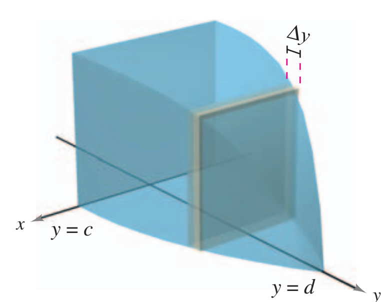
$$\text{Volume} = \int_a^b A(x)\,dx.$$ See Figure 7.24(a).

2. For cross sections of area $A(y)$ taken perpendicular to the y-axis,

$$\text{Volume} = \int_c^d A(y)\,dy.$$ See Figure 7.24(b).



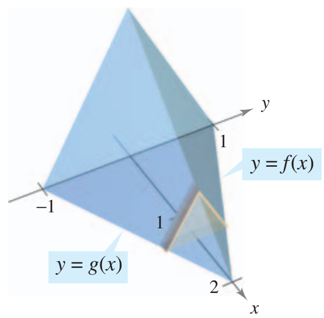
(a) Cross sections perpendicular to x-axis



(b) Cross sections perpendicular to y-axis

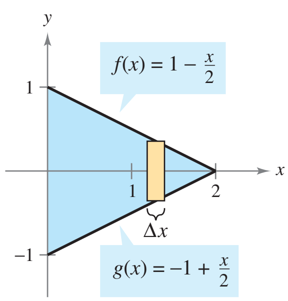
Figure 7.24

Rotatable Graph Rotatable Graph



Cross sections are equilateral triangles.

Rotatable Graph



Triangular base in xy-plane

Figure 7.25

EXAMPLE 6 Triangular Cross Sections

Find the volume of the solid shown in Figure 7.25. The base of the solid is the region bounded by the lines

$$f(x) = 1 - \frac{x}{2}, \qquad g(x) = -1 + \frac{x}{2}, \qquad \text{and} \qquad x = 0.$$

The cross sections perpendicular to the x-axis are equilateral triangles.

Solution The base and area of each triangular cross section are as follows.

$$\text{Base} = \left(1 - \frac{x}{2}\right) - \left(-1 + \frac{x}{2}\right) = 2 - x \qquad \text{Length of base}$$

$$\text{Area} = \frac{\sqrt{3}}{4}(\text{base})^2 \qquad \text{Area of equilateral triangle}$$

$$A(x) = \frac{\sqrt{3}}{4}(2 - x)^2 \qquad \text{Area of cross section}$$

Because x ranges from 0 to 2, the volume of the solid is

$$V = \int_a^b A(x)\,dx = \int_0^2 \frac{\sqrt{3}}{4}(2 - x)^2\,dx$$

$$= -\frac{\sqrt{3}}{4}\left[\frac{(2 - x)^3}{3}\right]_0^2 = \frac{2\sqrt{3}}{3}.$$

Try It **Exploration A**

EXAMPLE 7 An Application to Geometry

Prove that the volume of a pyramid with a square base is $V = \frac{1}{3}hB$, where h is the height of the pyramid and B is the area of the base.

Solution As shown in Figure 7.26, you can intersect the pyramid with a plane parallel to the base at height y to form a square cross section whose sides are of length b'. Using similar triangles, you can show that

$$\frac{b'}{b} = \frac{h - y}{h} \qquad \text{or} \qquad b' = \frac{b}{h}(h - y)$$

where b is the length of the sides of the base of the pyramid. So,

$$A(y) = (b')^2 = \frac{b^2}{h^2}(h - y)^2.$$

Integrating between 0 and h produces

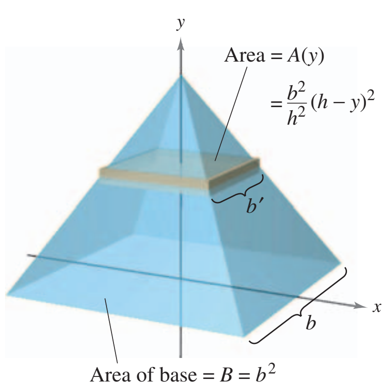
$$V = \int_0^h A(y)\,dy = \int_0^h \frac{b^2}{h^2}(h - y)^2\,dy$$

$$= \frac{b^2}{h^2}\int_0^h (h - y)^2\,dy$$

$$= -\left(\frac{b^2}{h^2}\right)\left[\frac{(h - y)^3}{3}\right]_0^h$$

$$= \frac{b^2}{h^2}\left(\frac{h^3}{3}\right)$$

$$= \frac{1}{3}hB. \qquad B = b^2$$



Rotatable Graph

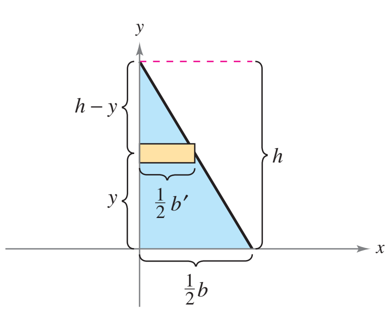


Figure 7.26

Try It **Exploration A** **Exploration B**

Section 7.3

Volume: The Shell Method

- **Find the volume of a solid of revolution using the shell method.**
- **Compare the uses of the disk method and the shell method.**

The Shell Method

In this section, you will study an alternative method for finding the volume of a solid of revolution. This method is called the **shell method** because it uses cylindrical shells. A comparison of the advantages of the disk and shell methods is given later in this section.

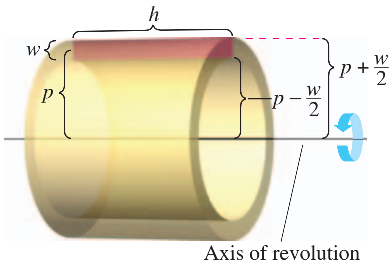


Figure 7.27

Rotatable Graph

To begin, consider a representative rectangle as shown in Figure 7.27, where w is the width of the rectangle, h is the height of the rectangle, and p is the distance between the axis of revolution and the *center* of the rectangle. When this rectangle is revolved about its axis of revolution, it forms a cylindrical shell (or tube) of thickness w. To find the volume of this shell, consider two cylinders. The radius of the larger cylinder corresponds to the outer radius of the shell, and the radius of the smaller cylinder corresponds to the inner radius of the shell. Because p is the average radius of the shell, you know the outer radius is $p + (w/2)$ and the inner radius is $p - (w/2)$.

$$p + \frac{w}{2} \qquad \text{Outer radius}$$

$$p - \frac{w}{2} \qquad \text{Inner radius}$$

So, the volume of the shell is

$$\begin{aligned}
\text{Volume of shell} &= (\text{volume of cylinder}) - (\text{volume of hole}) \\
&= \pi\left(p + \frac{w}{2}\right)^2 h - \pi\left(p - \frac{w}{2}\right)^2 h \\
&= 2\pi phw \\
&= 2\pi(\text{average radius})(\text{height})(\text{thickness}).
\end{aligned}$$

You can use this formula to find the volume of a solid of revolution. Assume that the plane region in Figure 7.28 is revolved about a line to form the indicated solid. If you consider a horizontal rectangle of width Δy, then, as the plane region is revolved about a line parallel to the x-axis, the rectangle generates a representative shell whose volume is

$$\Delta V = 2\pi[p(y)h(y)]\,\Delta y.$$

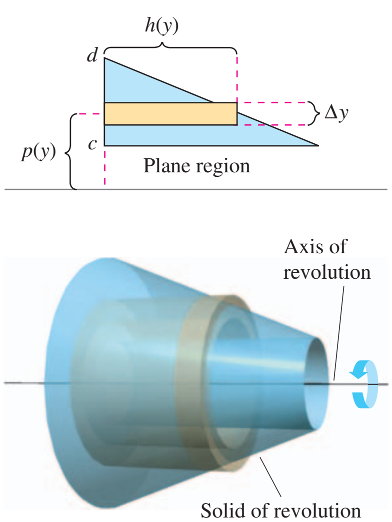


Figure 7.28

Rotatable Graph

You can approximate the volume of the solid by n such shells of thickness Δy, height $h(y_i)$, and average radius $p(y_i)$.

$$\text{Volume of solid} \approx \sum_{i=1}^{n} 2\pi[p(y_i)h(y_i)]\,\Delta y = 2\pi\sum_{i=1}^{n}[p(y_i)h(y_i)]\,\Delta y$$

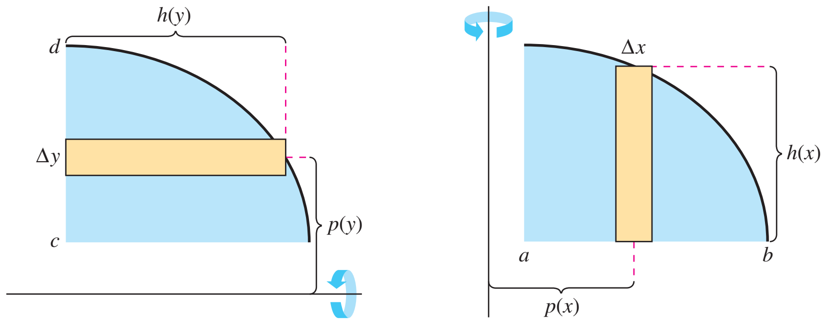
This approximation appears to become better and better as $\|\Delta\| \to 0$ $(n \to \infty)$. So, the volume of the solid is

$$\begin{aligned}
\text{Volume of solid} &= \lim_{\|\Delta\|\to 0} 2\pi\sum_{i=1}^{n}[p(y_i)h(y_i)]\,\Delta y \\
&= 2\pi\int_c^d [p(y)h(y)]\,dy.
\end{aligned}$$

The Shell Method

To find the volume of a solid of revolution with the **shell method,** use one of the following, as shown in Figure 7.29.

Horizontal Axis of Revolution	*Vertical Axis of Revolution*
$\text{Volume} = V = 2\pi \int_c^d p(y)h(y)\,dy$	$\text{Volume} = V = 2\pi \int_a^b p(x)h(x)\,dx$



Horizontal axis of revolution

Vertical axis of revolution

Figure 7.29

EXAMPLE 1 Using the Shell Method to Find Volume

Find the volume of the solid of revolution formed by revolving the region bounded by

$$y = x - x^3$$

and the x-axis $(0 \le x \le 1)$ about the y-axis.

Solution Because the axis of revolution is vertical, use a vertical representative rectangle, as shown in Figure 7.30. The width Δx indicates that x is the variable of integration. The distance from the center of the rectangle to the axis of revolution is $p(x) = x$, and the height of the rectangle is

$$h(x) = x - x^3.$$

Because x ranges from 0 to 1, the volume of the solid is

$$\begin{aligned} V = 2\pi \int_a^b p(x)h(x)\,dx &= 2\pi \int_0^1 x(x - x^3)\,dx && \text{Apply shell method.} \\ &= 2\pi \int_0^1 (-x^4 + x^2)\,dx && \text{Simplify.} \\ &= 2\pi \left[-\frac{x^5}{5} + \frac{x^3}{3} \right]_0^1 && \text{Integrate.} \\ &= 2\pi \left(-\frac{1}{5} + \frac{1}{3} \right) \\ &= \frac{4\pi}{15}. \end{aligned}$$

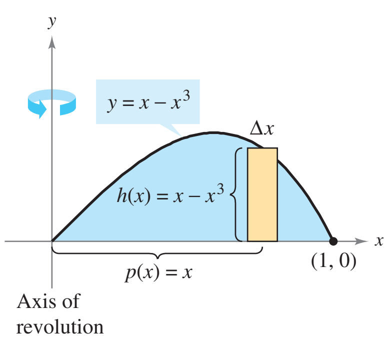


Figure 7.30

Try It

Exploration A

EXAMPLE 2 Using the Shell Method to Find Volume

Find the volume of the solid of revolution formed by revolving the region bounded by the graph of

$$x = e^{-y^2}$$

and the y-axis $(0 \le y \le 1)$ about the x-axis.

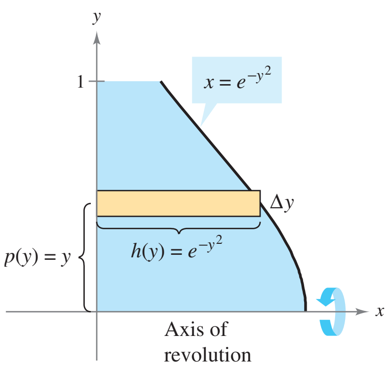


Figure 7.31

Solution Because the axis of revolution is horizontal, use a horizontal representative rectangle, as shown in Figure 7.31. The width Δy indicates that y is the variable of integration. The distance from the center of the rectangle to the axis of revolution is $p(y) = y$, and the height of the rectangle is $h(y) = e^{-y^2}$. Because y ranges from 0 to 1, the volume of the solid is

$$V = 2\pi \int_c^d p(y)h(y)\, dy = 2\pi \int_0^1 y e^{-y^2}\, dy \qquad \text{Apply shell method.}$$

$$= -\pi \Big[e^{-y^2} \Big]_0^1 \qquad \text{Integrate.}$$

$$= \pi\left(1 - \frac{1}{e}\right)$$

$$\approx 1.986.$$

Try It **Exploration A**

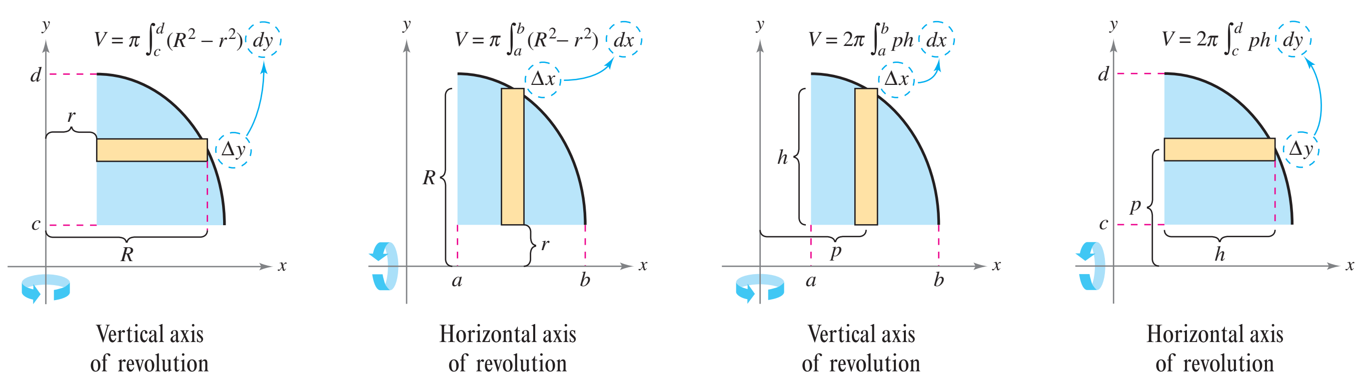
NOTE To see the advantage of using the shell method in Example 2, solve the equation $x = e^{-y^2}$ for y.

$$y = \begin{cases} 1, & 0 \le x \le 1/e \\ \sqrt{-\ln x}, & 1/e < x \le 1 \end{cases}$$

Then use this equation to find the volume using the disk method.

Comparison of Disk and Shell Methods

The disk and shell methods can be distinguished as follows. For the disk method, the representative rectangle is always *perpendicular* to the axis of revolution, whereas for the shell method, the representative rectangle is always *parallel* to the axis of revolution, as shown in Figure 7.32.



Disk method: Representative rectangle is perpendicular to the axis of revolution.

Shell method: Representative rectangle is parallel to the axis of revolution.

Figure 7.32

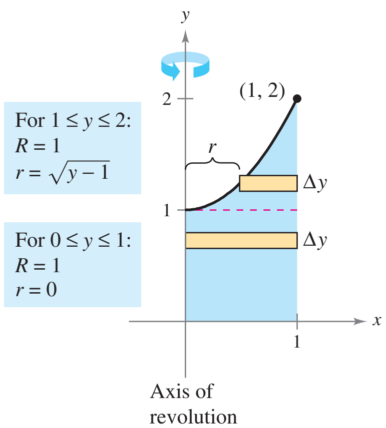
Often, one method is more convenient to use than the other. The following example illustrates a case in which the shell method is preferable.

EXAMPLE 3 Shell Method Preferable

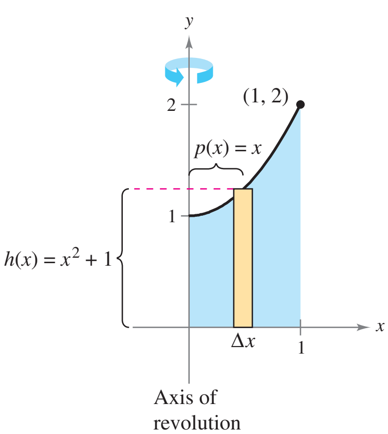
Find the volume of the solid formed by revolving the region bounded by the graphs of

$$y = x^2 + 1, \quad y = 0, \quad x = 0, \quad \text{and} \quad x = 1$$

about the y-axis.



(a) Disk method



(b) Shell method

Figure 7.33

Solution In Example 4 in the preceding section, you saw that the washer method requires two integrals to determine the volume of this solid. See Figure 7.33(a).

$$\begin{aligned}
V &= \pi \int_0^1 (1^2 - 0^2)\,dy + \pi \int_1^2 \left[1^2 - \left(\sqrt{y-1}\right)^2\right] dy && \text{Apply washer method.}\\
&= \pi \int_0^1 1\,dy + \pi \int_1^2 (2 - y)\,dy && \text{Simplify.}\\
&= \pi\Big[y\Big]_0^1 + \pi\left[2y - \frac{y^2}{2}\right]_1^2 && \text{Integrate.}\\
&= \pi + \pi\left(4 - 2 - 2 + \frac{1}{2}\right)\\
&= \frac{3\pi}{2}
\end{aligned}$$

In Figure 7.33(b), you can see that the shell method requires only one integral to find the volume.

$$\begin{aligned}
V &= 2\pi \int_a^b p(x)h(x)\,dx && \text{Apply shell method.}\\
&= 2\pi \int_0^1 x(x^2 + 1)\,dx\\
&= 2\pi \left[\frac{x^4}{4} + \frac{x^2}{2}\right]_0^1 && \text{Integrate.}\\
&= 2\pi\left(\frac{3}{4}\right)\\
&= \frac{3\pi}{2}
\end{aligned}$$

Try It | Exploration A | Open Exploration

Suppose the region in Example 3 were revolved about the vertical line $x = 1$. Would the resulting solid of revolution have a greater volume or a smaller volume than the solid in Example 3? Without integrating, you should be able to reason that the resulting solid would have a smaller volume because "more" of the revolved region would be closer to the axis of revolution. To confirm this, try solving the following integral, which gives the volume of the solid.

$$V = 2\pi \int_0^1 (1 - x)(x^2 + 1)\,dx \qquad p(x) = 1 - x$$

FOR FURTHER INFORMATION To learn more about the disk and shell methods, see the article "The Disk and Shell Method" by Charles A. Cable in *The American Mathematical Monthly*.

MathArticle

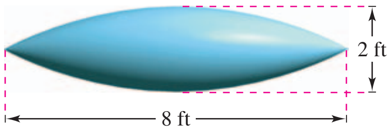
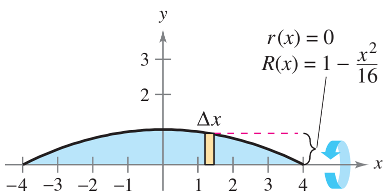
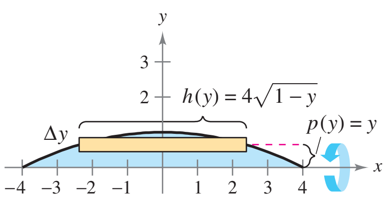


Figure 7.34

Rotatable Graph



(a) Disk method



(b) Shell method

Figure 7.35

EXAMPLE 4 Volume of a Pontoon

A pontoon is to be made in the shape shown in Figure 7.34. The pontoon is designed by rotating the graph of

$$y = 1 - \frac{x^2}{16}, \qquad -4 \le x \le 4$$

about the x-axis, where x and y are measured in feet. Find the volume of the pontoon.

Solution Refer to Figure 7.35(a) and use the disk method as follows.

$$\begin{aligned} V &= \pi \int_{-4}^{4} \left(1 - \frac{x^2}{16}\right)^2 dx && \text{Apply disk method.} \\ &= \pi \int_{-4}^{4} \left(1 - \frac{x^2}{8} + \frac{x^4}{256}\right) dx && \text{Simplify.} \\ &= \pi \left[x - \frac{x^3}{24} + \frac{x^5}{1280} \right]_{-4}^{4} && \text{Integrate.} \\ &= \frac{64\pi}{15} \approx 13.4 \text{ cubic feet} \end{aligned}$$

Try using Figure 7.35(b) to set up the integral for the volume using the shell method. Does the integral seem more complicated?

Try It Exploration A

For the shell method in Example 4, you would have to solve for x in terms of y in the equation

$$y = 1 - (x^2/16).$$

Sometimes, solving for x is very difficult (or even impossible). In such cases you must use a vertical rectangle (of width Δx), thus making x the variable of integration. The position (horizontal or vertical) of the axis of revolution then determines the method to be used. This is shown in Example 5.

EXAMPLE 5 Shell Method Necessary

Find the volume of the solid formed by revolving the region bounded by the graphs of $y = x^3 + x + 1$, $y = 1$, and $x = 1$ about the line $x = 2$, as shown in Figure 7.36.

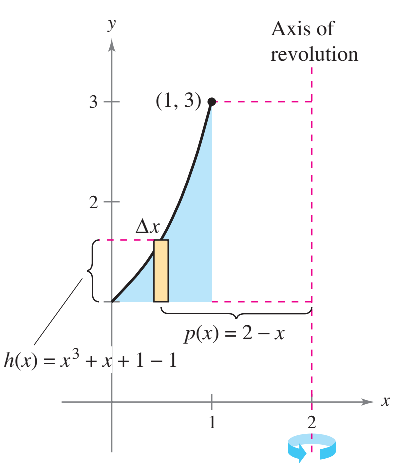


Figure 7.36

Solution In the equation $y = x^3 + x + 1$, you cannot easily solve for x in terms of y. (See Section 3.8 on Newton's Method.) Therefore, the variable of integration must be x, and you should choose a vertical representative rectangle. Because the rectangle is parallel to the axis of revolution, use the shell method and obtain

$$\begin{aligned} V = 2\pi \int_a^b p(x)h(x)\,dx &= 2\pi \int_0^1 (2 - x)(x^3 + x + 1 - 1)\,dx && \text{Apply shell method.} \\ &= 2\pi \int_0^1 (-x^4 + 2x^3 - x^2 + 2x)\,dx && \text{Simplify.} \\ &= 2\pi \left[-\frac{x^5}{5} + \frac{x^4}{2} - \frac{x^3}{3} + x^2 \right]_0^1 && \text{Integrate.} \\ &= 2\pi \left(-\frac{1}{5} + \frac{1}{2} - \frac{1}{3} + 1 \right) \\ &= \frac{29\pi}{15}. \end{aligned}$$

Try It Exploration A

Section 7.4

Arc Length and Surfaces of Revolution

- **Find the arc length of a smooth curve.**
- **Find the area of a surface of revolution.**

Christian Huygens (1629–1695)

The Dutch mathematician Christian Huygens, who invented the pendulum clock, and James Gregory (1638–1675), a Scottish mathematician, both made early contributions to the problem of finding the length of a rectifiable curve.

MathBio

History

Arc Length

In this section, definite integrals are used to find the arc lengths of curves and the areas of surfaces of revolution. In either case, an arc (a segment of a curve) is approximated by straight line segments whose lengths are given by the familiar Distance Formula

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}.$$

A **rectifiable** curve is one that has a finite arc length. You will see that a sufficient condition for the graph of a function f to be rectifiable between $(a, f(a))$ and $(b, f(b))$ is that f' be continuous on $[a, b]$. Such a function is **continuously differentiable** on $[a, b]$, and its graph on the interval $[a, b]$ is a **smooth curve.**

Consider a function $y = f(x)$ that is continuously differentiable on the interval $[a, b]$. You can approximate the graph of f by n line segments whose endpoints are determined by the partition

$$a = x_0 < x_1 < x_2 < \cdot \cdot \cdot < x_n = b$$

as shown in Figure 7.37. By letting $\Delta x_i = x_i - x_{i-1}$ and $\Delta y_i = y_i - y_{i-1}$, you can approximate the length of the graph by

$$\begin{aligned} s &\approx \sum_{i=1}^{n} \sqrt{(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2} \\ &= \sum_{i=1}^{n} \sqrt{(\Delta x_i)^2 + (\Delta y_i)^2} \\ &= \sum_{i=1}^{n} \sqrt{(\Delta x_i)^2 + \left(\frac{\Delta y_i}{\Delta x_i}\right)^2 (\Delta x_i)^2} \\ &= \sum_{i=1}^{n} \sqrt{1 + \left(\frac{\Delta y_i}{\Delta x_i}\right)^2}\,(\Delta x_i). \end{aligned}$$

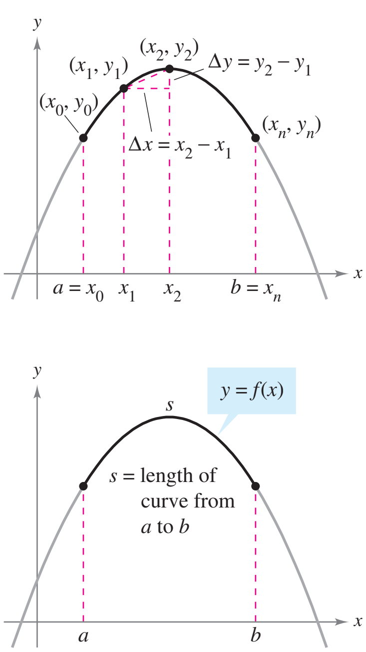


Figure 7.37

This approximation appears to become better and better as $\|\Delta\| \to 0$ $(n \to \infty)$. So, the length of the graph is

$$s = \lim_{\|\Delta\| \to 0} \sum_{i=1}^{n} \sqrt{1 + \left(\frac{\Delta y_i}{\Delta x_i}\right)^2}\,(\Delta x_i).$$

Because $f'(x)$ exists for each x in (x_{i-1}, x_i), the Mean Value Theorem guarantees the existence of c_i in (x_{i-1}, x_i) such that

$$\begin{aligned} f(x_i) - f(x_{i-1}) &= f'(c_i)(x_i - x_{i-1}) \\ \frac{\Delta y_i}{\Delta x_i} &= f'(c_i). \end{aligned}$$

Because f' is continuous on $[a, b]$, it follows that $\sqrt{1 + [f'(x)]^2}$ is also continuous (and therefore integrable) on $[a, b]$, which implies that

$$\begin{aligned} s &= \lim_{\|\Delta\| \to 0} \sum_{i=1}^{n} \sqrt{1 + [f'(c_i)]^2}\,(\Delta x_i) \\ &= \int_a^b \sqrt{1 + [f'(x)]^2}\, dx \end{aligned}$$

where s is called the **arc length** of f between a and b.

Definition of Arc Length

Let the function given by $y = f(x)$ represent a smooth curve on the interval $[a, b]$. The **arc length** of f between a and b is

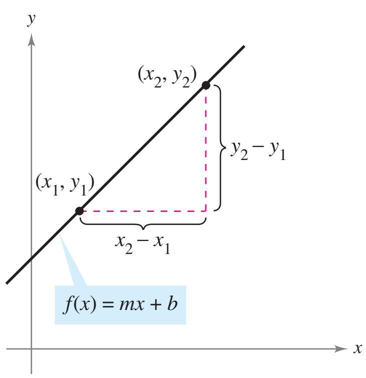
$$s = \int_a^b \sqrt{1 + [f'(x)]^2}\, dx.$$

Similarly, for a smooth curve given by $x = g(y)$, the **arc length** of g between c and d is

$$s = \int_c^d \sqrt{1 + [g'(y)]^2}\, dy.$$

Because the definition of arc length can be applied to a linear function, you can check to see that this new definition agrees with the standard Distance Formula for the length of a line segment. This is shown in Example 1.

Technology



The arc length of the graph of f from (x_1, y_1) to (x_2, y_2) is the same as the standard Distance Formula.
Figure 7.38

EXAMPLE 1 The Length of a Line Segment

Find the arc length from (x_1, y_1) to (x_2, y_2) on the graph of $f(x) = mx + b$, as shown in Figure 7.38.

Solution Because

$$m = f'(x) = \frac{y_2 - y_1}{x_2 - x_1}$$

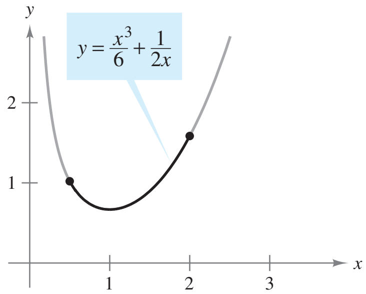
it follows that

$$\begin{aligned}
s &= \int_{x_1}^{x_2} \sqrt{1 + [f'(x)]^2}\, dx && \text{Formula for arc length} \\
&= \int_{x_1}^{x_2} \sqrt{1 + \left(\frac{y_2 - y_1}{x_2 - x_1}\right)^2}\, dx \\
&= \sqrt{\frac{(x_2 - x_1)^2 + (y_2 - y_1)^2}{(x_2 - x_1)^2}}\,(x)\Big]_{x_1}^{x_2} && \text{Integrate and simplify.} \\
&= \sqrt{\frac{(x_2 - x_1)^2 + (y_2 - y_1)^2}{(x_2 - x_1)^2}}\,(x_2 - x_1) \\
&= \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}
\end{aligned}$$

which is the formula for the distance between two points in the plane.

Try It   Exploration A

TECHNOLOGY Definite integrals representing arc length often are very difficult to evaluate. In this section, a few examples are presented. In the next chapter, with more advanced integration techniques, you will be able to tackle more difficult arc length problems. In the meantime, remember that you can always use a numerical integration program to approximate an arc length. For instance, use the *numerical integration* feature of a graphing utility to approximate the arc lengths in Examples 2 and 3.



The arc length of the graph of y on $\left[\frac{1}{2}, 2\right]$
Figure 7.39

Editable Graph

FOR FURTHER INFORMATION To see how arc length can be used to define trigonometric functions, see the article "Trigonometry Requires Calculus, Not Vice Versa" by Yves Nievergelt in *UMAP Modules*.

EXAMPLE 2 Finding Arc Length

Find the arc length of the graph of

$$y = \frac{x^3}{6} + \frac{1}{2x}$$

on the interval $\left[\frac{1}{2}, 2\right]$, as shown in Figure 7.39.

Solution Using

$$\frac{dy}{dx} = \frac{3x^2}{6} - \frac{1}{2x^2} = \frac{1}{2}\left(x^2 - \frac{1}{x^2}\right)$$

yields an arc length of

$$\begin{aligned} s = \int_a^b \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\, dx &= \int_{1/2}^{2} \sqrt{1 + \left[\frac{1}{2}\left(x^2 - \frac{1}{x^2}\right)\right]^2}\, dx && \text{Formula for arc length} \\ &= \int_{1/2}^{2} \sqrt{\frac{1}{4}\left(x^4 + 2 + \frac{1}{x^4}\right)}\, dx \\ &= \int_{1/2}^{2} \frac{1}{2}\left(x^2 + \frac{1}{x^2}\right) dx && \text{Simplify.} \\ &= \frac{1}{2}\left[\frac{x^3}{3} - \frac{1}{x}\right]_{1/2}^{2} && \text{Integrate.} \\ &= \frac{1}{2}\left(\frac{13}{6} + \frac{47}{24}\right) \\ &= \frac{33}{16}. \end{aligned}$$

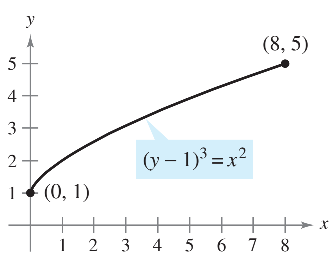
Try It Exploration A

EXAMPLE 3 Finding Arc Length

Find the arc length of the graph of $(y - 1)^3 = x^2$ on the interval $[0, 8]$, as shown in Figure 7.40.

Solution Begin by solving for x in terms of y: $x = \pm(y - 1)^{3/2}$. Choosing the positive value of x produces

$$\frac{dx}{dy} = \frac{3}{2}(y - 1)^{1/2}.$$



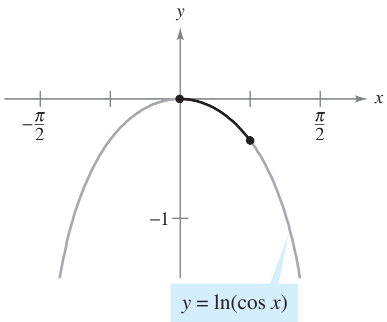
The arc length of the graph of y on $[0, 8]$
Figure 7.40

Editable Graph

The x-interval $[0, 8]$ corresponds to the y-interval $[1, 5]$, and the arc length is

$$\begin{aligned} s = \int_c^d \sqrt{1 + \left(\frac{dx}{dy}\right)^2}\, dy &= \int_1^5 \sqrt{1 + \left[\frac{3}{2}(y - 1)^{1/2}\right]^2}\, dy && \text{Formula for arc length} \\ &= \int_1^5 \sqrt{\frac{9}{4}y - \frac{5}{4}}\, dy \\ &= \frac{1}{2}\int_1^5 \sqrt{9y - 5}\, dy && \text{Simplify.} \\ &= \frac{1}{18}\left[\frac{(9y - 5)^{3/2}}{3/2}\right]_1^5 && \text{Integrate.} \\ &= \frac{1}{27}(40^{3/2} - 4^{3/2}) \\ &\approx 9.073. \end{aligned}$$

Try It Exploration A



The arc length of the graph of y on $\left[0, \frac{\pi}{4}\right]$

Figure 7.41

Editable Graph

EXAMPLE 4 Finding Arc Length

Find the arc length of the graph of $y = \ln(\cos x)$ from $x = 0$ to $x = \pi/4$, as shown in Figure 7.41.

Solution Using

$$\frac{dy}{dx} = -\frac{\sin x}{\cos x} = -\tan x$$

yields an arc length of

$$\begin{aligned}
s = \int_a^b \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\,dx &= \int_0^{\pi/4} \sqrt{1 + \tan^2 x}\,dx && \text{Formula for arc length}\\
&= \int_0^{\pi/4} \sqrt{\sec^2 x}\,dx && \text{Trigonometric identity}\\
&= \int_0^{\pi/4} \sec x\,dx && \text{Simplify.}\\
&= \Big[\ln|\sec x + \tan x|\Big]_0^{\pi/4} && \text{Integrate.}\\
&= \ln\left(\sqrt{2} + 1\right) - \ln 1\\
&\approx 0.881.
\end{aligned}$$

Try It Exploration A Open Exploration

EXAMPLE 5 Length of a Cable

An electric cable is hung between two towers that are 200 feet apart, as shown in Figure 7.42. The cable takes the shape of a catenary whose equation is

$$y = 75\left(e^{x/150} + e^{-x/150}\right) = 150 \cosh \frac{x}{150}.$$

Find the arc length of the cable between the two towers.

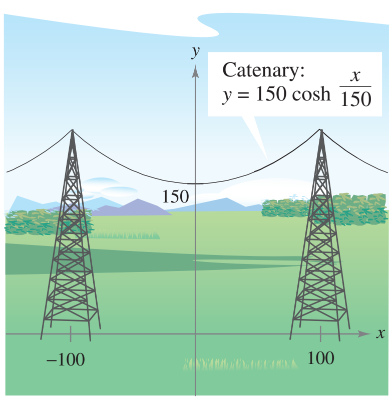


Figure 7.42

Solution Because $y' = \frac{1}{2}\left(e^{x/150} - e^{-x/150}\right)$, you can write

$$(y')^2 = \frac{1}{4}\left(e^{x/75} - 2 + e^{-x/75}\right)$$

and

$$1 + (y')^2 = \frac{1}{4}\left(e^{x/75} + 2 + e^{-x/75}\right) = \left[\frac{1}{2}\left(e^{x/150} + e^{-x/150}\right)\right]^2.$$

Therefore, the arc length of the cable is

$$\begin{aligned}
s = \int_a^b \sqrt{1 + (y')^2}\,dx &= \frac{1}{2}\int_{-100}^{100} \left(e^{x/150} + e^{-x/150}\right) dx && \text{Formula for arc length}\\
&= 75\Big[e^{x/150} - e^{-x/150}\Big]_{-100}^{100} && \text{Integrate.}\\
&= 150\left(e^{2/3} - e^{-2/3}\right)\\
&\approx 215 \text{ feet.}
\end{aligned}$$

Try It Exploration A

Area of a Surface of Revolution

In Sections 7.2 and 7.3, integration was used to calculate the volume of a solid of revolution. You will now look at a procedure for finding the area of a surface of revolution.

> **Definition of Surface of Revolution**
>
> If the graph of a continuous function is revolved about a line, the resulting surface is a **surface of revolution.**

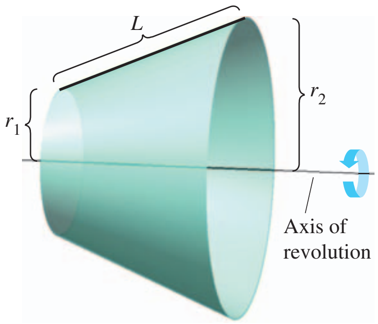


Figure 7.43

Rotatable Graph

The area of a surface of revolution is derived from the formula for the lateral surface area of the frustum of a right circular cone. Consider the line segment in Figure 7.43, where L is the length of the line segment, r_1 is the radius at the left end of the line segment, and r_2 is the radius at the right end of the line segment. When the line segment is revolved about its axis of revolution, it forms a frustum of a right circular cone, with

$$S = 2\pi r L \qquad \text{Lateral surface area of frustum}$$

where

$$r = \frac{1}{2}(r_1 + r_2). \qquad \text{Average radius of frustum}$$

(In Exercise 60, you are asked to verify the formula for S.)

Suppose the graph of a function f, having a continuous derivative on the interval $[a, b]$, is revolved about the x-axis to form a surface of revolution, as shown in Figure 7.44. Let Δ be a partition of $[a, b]$, with subintervals of width Δx_i. Then the line segment of length

$$\Delta L_i = \sqrt{\Delta x_i^2 + \Delta y_i^2}$$

generates a frustum of a cone. Let r_i be the average radius of this frustum. By the Intermediate Value Theorem, a point d_i exists (in the ith subinterval) such that $r_i = f(d_i)$. The lateral surface area ΔS_i of the frustum is

$$\begin{aligned}\Delta S_i &= 2\pi r_i \, \Delta L_i \\ &= 2\pi f(d_i)\sqrt{\Delta x_i^2 + \Delta y_i^2} \\ &= 2\pi f(d_i)\sqrt{1 + \left(\frac{\Delta y_i}{\Delta x_i}\right)^2}\, \Delta x_i.\end{aligned}$$

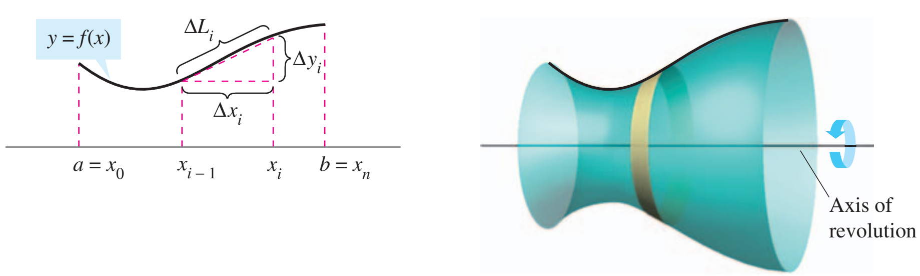
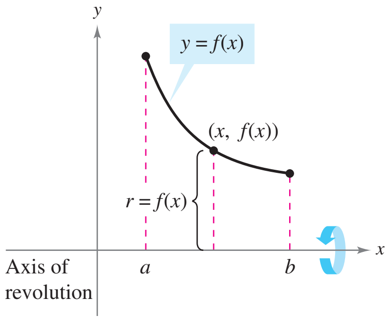


Figure 7.44

Rotatable Graph



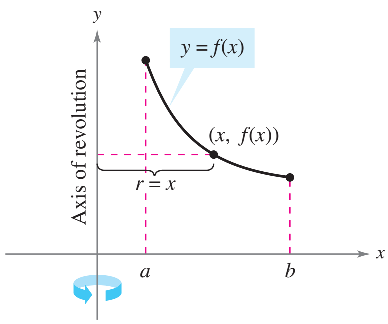


Figure 7.45

By the Mean Value Theorem, a point c_i exists in (x_{i-1}, x_i) such that

$$f'(c_i) = \frac{f(x_i) - f(x_{i-1})}{x_i - x_{i-1}}$$

$$= \frac{\Delta y_i}{\Delta x_i}.$$

So, $\Delta S_i = 2\pi f(d_i)\sqrt{1 + [f'(c_i)]^2}\,\Delta x_i$, and the total surface area can be approximated by

$$S \approx 2\pi \sum_{i=1}^{n} f(d_i)\sqrt{1 + [f'(c_i)]^2}\,\Delta x_i.$$

It can be shown that the limit of the right side as $\|\Delta\| \to 0$ $(n \to \infty)$ is

$$S = 2\pi \int_a^b f(x)\sqrt{1 + [f'(x)]^2}\,dx.$$

In a similar manner, if the graph of f is revolved about the y-axis, then S is

$$S = 2\pi \int_a^b x\sqrt{1 + [f'(x)]^2}\,dx.$$

In both formulas for S, you can regard the products $2\pi f(x)$ and $2\pi x$ as the circumference of the circle traced by a point (x, y) on the graph of f as it is revolved about the x- or y-axis (Figure 7.45). In one case the radius is $r = f(x)$, and in the other case the radius is $r = x$. Moreover, by appropriately adjusting r, you can generalize the formula for surface area to cover *any* horizontal or vertical axis of revolution, as indicated in the following definition.

Definition of the Area of a Surface of Revolution

Let $y = f(x)$ have a continuous derivative on the interval $[a, b]$. The area S of the surface of revolution formed by revolving the graph of f about a horizontal or vertical axis is

$$S = 2\pi \int_a^b r(x)\sqrt{1 + [f'(x)]^2}\,dx \qquad y \text{ is a function of } x.$$

where $r(x)$ is the distance between the graph of f and the axis of revolution. If $x = g(y)$ on the interval $[c, d]$, then the surface area is

$$S = 2\pi \int_c^d r(y)\sqrt{1 + [g'(y)]^2}\,dy \qquad x \text{ is a function of } y.$$

where $r(y)$ is the distance between the graph of g and the axis of revolution.

The formulas in this definition are sometimes written as

$$S = 2\pi \int_a^b r(x)\,ds \qquad y \text{ is a function of } x.$$

and

$$S = 2\pi \int_c^d r(y)\,ds \qquad x \text{ is a function of } y.$$

where $ds = \sqrt{1 + [f'(x)]^2}\,dx$ and $ds = \sqrt{1 + [g'(y)]^2}\,dy$, respectively.

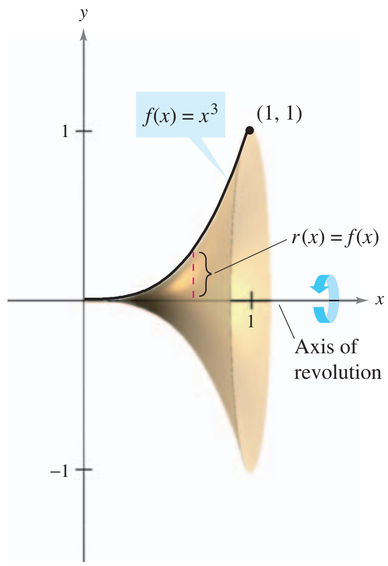


Figure 7.46

Rotatable Graph

EXAMPLE 6 The Area of a Surface of Revolution

Find the area of the surface formed by revolving the graph of

$$f(x) = x^3$$

on the interval $[0, 1]$ about the x-axis, as shown in Figure 7.46.

Solution The distance between the x-axis and the graph of f is $r(x) = f(x)$, and because $f'(x) = 3x^2$, the surface area is

$$\begin{aligned}
S &= 2\pi\int_a^b r(x)\sqrt{1 + [f'(x)]^2}\,dx && \text{Formula for surface area}\\
&= 2\pi\int_0^1 x^3\sqrt{1 + (3x^2)^2}\,dx\\
&= \frac{2\pi}{36}\int_0^1 (36x^3)(1 + 9x^4)^{1/2}\,dx && \text{Simplify.}\\
&= \frac{\pi}{18}\left[\frac{(1 + 9x^4)^{3/2}}{3/2}\right]_0^1 && \text{Integrate.}\\
&= \frac{\pi}{27}(10^{3/2} - 1)\\
&\approx 3.563.
\end{aligned}$$

Try It Exploration A

EXAMPLE 7 The Area of a Surface of Revolution

Find the area of the surface formed by revolving the graph of

$$f(x) = x^2$$

on the interval $[0, \sqrt{2}]$ about the y-axis, as shown in Figure 7.47.

Solution In this case, the distance between the graph of f and the y-axis is $r(x) = x$. Using $f'(x) = 2x$, you can determine that the surface area is

$$\begin{aligned}
S &= 2\pi\int_a^b r(x)\sqrt{1 + [f'(x)]^2}\,dx && \text{Formula for surface area}\\
&= 2\pi\int_0^{\sqrt{2}} x\sqrt{1 + (2x)^2}\,dx\\
&= \frac{2\pi}{8}\int_0^{\sqrt{2}} (1 + 4x^2)^{1/2}(8x)\,dx && \text{Simplify.}\\
&= \frac{\pi}{4}\left[\frac{(1 + 4x^2)^{3/2}}{3/2}\right]_0^{\sqrt{2}} && \text{Integrate.}\\
&= \frac{\pi}{6}[(1 + 8)^{3/2} - 1]\\
&= \frac{13\pi}{3}\\
&\approx 13.614.
\end{aligned}$$

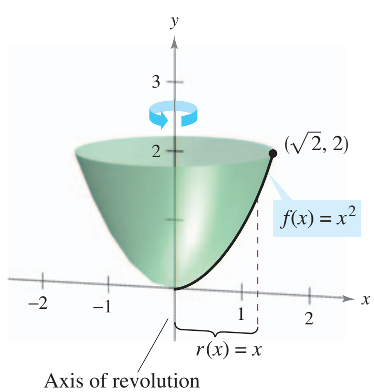


Figure 7.47

Rotatable Graph

Try It Exploration A

Section 7.5 Work

- **Find the work done by a constant force.**
- **Find the work done by a variable force.**

Work Done by a Constant Force

The concept of work is important to scientists and engineers for determining the energy needed to perform various jobs. For instance, it is useful to know the amount of work done when a crane lifts a steel girder, when a spring is compressed, when a rocket is propelled into the air, or when a truck pulls a load along a highway.

In general, **work** is done by a force when it moves an object. If the force applied to the object is *constant*, then the definition of work is as follows.

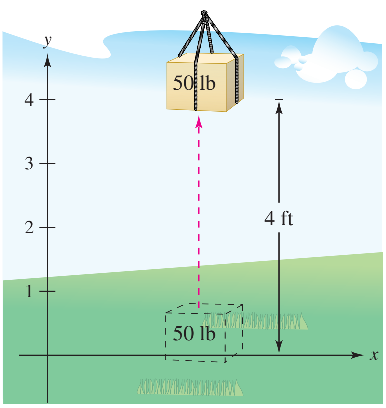
Definition of Work Done by a Constant Force

If an object is moved a distance D in the direction of an applied constant force F, then the **work** W done by the force is defined as $W = FD$.

There are many types of forces—centrifugal, electromotive, and gravitational, to name a few. A **force** can be thought of as a *push* or a *pull*; a force changes the state of rest or state of motion of a body. For gravitational forces on Earth, it is common to use units of measure corresponding to the weight of an object.

EXAMPLE 1 Lifting an Object

Determine the work done in lifting a 50-pound object 4 feet.



The work done in lifting a 50-pound object 4 feet is 200 foot-pounds.
Figure 7.48

Simulation

Solution The magnitude of the required force F is the weight of the object, as shown in Figure 7.48. So, the work done in lifting the object 4 feet is

$$
\begin{aligned}
W &= FD && \text{Work} = \text{(force)(distance)} \\
&= 50(4) && \text{Force} = 50 \text{ pounds, distance} = 4 \text{ feet} \\
&= 200 \text{ foot-pounds.}
\end{aligned}
$$

Try It | Exploration A | Exploration B

In the U.S. measurement system, work is typically expressed in foot-pounds (ft-lb), inch-pounds, or foot-tons. In the centimeter-gram-second (C-G-S) system, the basic unit of force is the **dyne**—the force required to produce an acceleration of 1 centimeter per second per second on a mass of 1 gram. In this system, work is typically expressed in dyne-centimeters (ergs) or newton-meters (joules), where 1 joule $= 10^7$ ergs.

EXPLORATION

How Much Work? In Example 1, 200 foot-pounds of work was needed to lift the 50-pound object 4 feet vertically off the ground. Suppose that once you lifted the object, you held it and walked a horizontal distance of 4 feet. Would this require an additional 200 foot-pounds of work? Explain your reasoning.

Work Done by a Variable Force

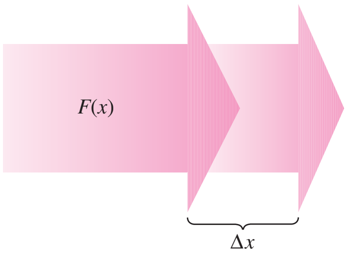
In Example 1, the force involved was *constant*. If a *variable* force is applied to an object, calculus is needed to determine the work done, because the amount of force changes as the object changes position. For instance, the force required to compress a spring increases as the spring is compressed.

Suppose that an object is moved along a straight line from $x = a$ to $x = b$ by a continuously varying force $F(x)$. Let Δ be a partition that divides the interval $[a, b]$ into n subintervals determined by

$$a = x_0 < x_1 < x_2 < \cdot \cdot \cdot < x_n = b$$

and let $\Delta x_i = x_i - x_{i-1}$. For each i, choose c_i such that $x_{i-1} \le c_i \le x_i$. Then at c_i the force is given by $F(c_i)$. Because F is continuous, you can approximate the work done in moving the object through the ith subinterval by the increment

$$\Delta W_i = F(c_i)\,\Delta x_i$$



The amount of force changes as an object changes position (Δx).
Figure 7.49

as shown in Figure 7.49. So, the total work done as the object moves from a to b is approximated by

$$\begin{aligned} W &\approx \sum_{i=1}^{n} \Delta W_i \\ &= \sum_{i=1}^{n} F(c_i)\,\Delta x_i. \end{aligned}$$

This approximation appears to become better and better as $\|\Delta\| \to 0$ $(n \to \infty)$. So, the work done is

$$\begin{aligned} W &= \lim_{\|\Delta\| \to 0} \sum_{i=1}^{n} F(c_i)\,\Delta x_i \\ &= \int_a^b F(x)\,dx. \end{aligned}$$

Definition of Work Done by a Variable Force

If an object is moved along a straight line by a continuously varying force $F(x)$, then the **work** W done by the force as the object is moved from $x = a$ to $x = b$ is

$$\begin{aligned} W &= \lim_{\|\Delta\| \to 0} \sum_{i=1}^{n} \Delta W_i \\ &= \int_a^b F(x)\,dx. \end{aligned}$$

EMILIE DE BRETEUIL (1706–1749)

Another major work by de Breteuil was the translation of Newton's "Philosophiae Naturalis Principia Mathematica" into French. Her translation and commentary greatly contributed to the acceptance of Newtonian science in Europe.

MathBio

The remaining examples in this section use some well-known physical laws. The discoveries of many of these laws occurred during the same period in which calculus was being developed. In fact, during the seventeenth and eighteenth centuries, there was little difference between physicists and mathematicians. One such physicist-mathematician was Emilie de Breteuil. Breteuil was instrumental in synthesizing the work of many other scientists, including Newton, Leibniz, Huygens, Kepler, and Descartes. Her physics text *Institutions* was widely used for many years.

The following three laws of physics were developed by Robert Hooke (1635–1703), Isaac Newton (1642–1727), and Charles Coulomb (1736–1806).

1. **Hooke's Law:** The force F required to compress or stretch a spring (within its elastic limits) is proportional to the distance d that the spring is compressed or stretched from its original length. That is,

$$F = kd$$

where the constant of proportionality k (the spring constant) depends on the specific nature of the spring.

2. **Newton's Law of Universal Gravitation:** The force F of attraction between two particles of masses m_1 and m_2 is proportional to the product of the masses and inversely proportional to the square of the distance d between the two particles. That is,

$$F = k\frac{m_1 m_2}{d^2}.$$

If m_1 and m_2 are given in grams and d in centimeters, F will be in dynes for a value of $k = 6.670 \times 10^{-8}$ cubic centimeter per gram-second squared.

3. **Coulomb's Law:** The force between two charges q_1 and q_2 in a vacuum is proportional to the product of the charges and inversely proportional to the square of the distance d between the two charges. That is,

$$F = k\frac{q_1 q_2}{d^2}.$$

If q_1 and q_2 are given in electrostatic units and d in centimeters, F will be in dynes for a value of $k = 1$.

EXPLORATION

The work done in compressing the spring in Example 2 from $x = 3$ inches to $x = 6$ inches is 3375 inch-pounds. Should the work done in compressing the spring from $x = 0$ inches to $x = 3$ inches be more than, the same as, or less than this? Explain.

EXAMPLE 2 Compressing a Spring

A force of 750 pounds compresses a spring 3 inches from its natural length of 15 inches. Find the work done in compressing the spring an additional 3 inches.

Solution By Hooke's Law, the force $F(x)$ required to compress the spring x units (from its natural length) is $F(x) = kx$. Using the given data, it follows that $F(3) = 750 = (k)(3)$ and so $k = 250$ and $F(x) = 250x$, as shown in Figure 7.50. To find the increment of work, assume that the force required to compress the spring over a small increment Δx is nearly constant. So, the increment of work is

$$\Delta W = (\text{force})(\text{distance increment}) = (250x)\,\Delta x.$$

Because the spring is compressed from $x = 3$ to $x = 6$ inches less than its natural length, the work required is

$$W = \int_a^b F(x)\,dx = \int_3^6 250x\,dx \qquad \text{Formula for work}$$

$$= 125x^2\Big]_3^6 = 4500 - 1125 = 3375 \text{ inch-pounds.}$$

Note that you do *not* integrate from $x = 0$ to $x = 6$ because you were asked to determine the work done in compressing the spring an *additional* 3 inches (not including the first 3 inches).

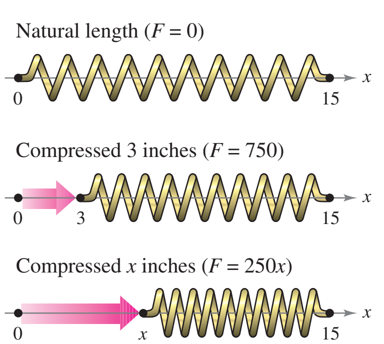


Figure 7.50

Try It | Exploration A | Exploration B | Open Exploration

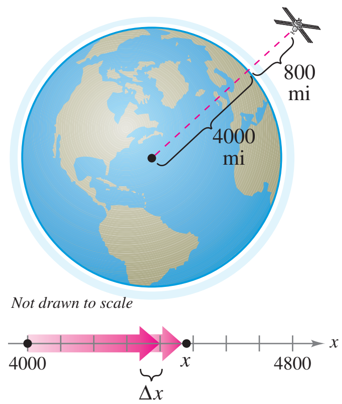


Figure 7.51

EXAMPLE 3 Moving a Space Module into Orbit

A space module weighs 15 metric tons on the surface of Earth. How much work is done in propelling the module to a height of 800 miles above Earth, as shown in Figure 7.51? (Use 4000 miles as the radius of Earth. Do not consider the effect of air resistance or the weight of the propellant.)

Solution Because the weight of a body varies inversely as the square of its distance from the center of Earth, the force $F(x)$ exerted by gravity is

$$F(x) = \frac{C}{x^2}. \qquad C \text{ is the constant of proportionality.}$$

Because the module weighs 15 metric tons on the surface of Earth and the radius of Earth is approximately 4000 miles, you have

$$\begin{aligned} 15 &= \frac{C}{(4000)^2} \\ 240{,}000{,}000 &= C. \end{aligned}$$

So, the increment of work is

$$\begin{aligned} \Delta W &= (\text{force})(\text{distance increment}) \\ &= \frac{240{,}000{,}000}{x^2}\,\Delta x. \end{aligned}$$

Finally, because the module is propelled from $x = 4000$ to $x = 4800$ miles, the total work done is

$$\begin{aligned} W &= \int_a^b F(x)\,dx = \int_{4000}^{4800} \frac{240{,}000{,}000}{x^2}\,dx && \text{Formula for work} \\ &= \frac{-240{,}000{,}000}{x}\Big]_{4000}^{4800} && \text{Integrate.} \\ &= -50{,}000 + 60{,}000 \\ &= 10{,}000 \text{ mile-tons} \\ &\approx 1.164 \times 10^{11} \text{ foot-pounds.} \end{aligned}$$

In the C-G-S system, using a conversion factor of 1 foot-pound $\approx$ 1.35582 joules, the work done is

$$W \approx 1.578 \times 10^{11} \text{ joules.}$$

Try It Exploration A Exploration B

The solutions to Examples 2 and 3 conform to our development of work as the summation of increments in the form

$$\Delta W = (\text{force})(\text{distance increment}) = (F)(\Delta x).$$

Another way to formulate the increment of work is

$$\Delta W = (\text{force increment})(\text{distance}) = (\Delta F)(x).$$

This second interpretation of ΔW is useful in problems involving the movement of nonrigid substances such as fluids and chains.

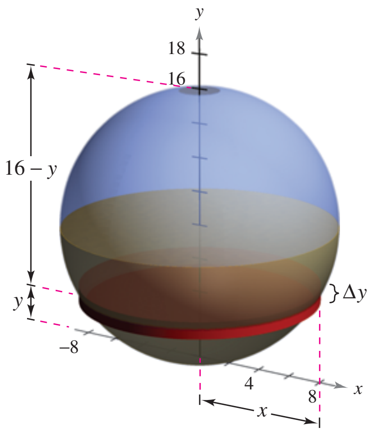


Figure 7.52

Rotatable Graph

EXAMPLE 4 Emptying a Tank of Oil

A spherical tank of radius 8 feet is half full of oil that weighs 50 pounds per cubic foot. Find the work required to pump oil out through a hole in the top of the tank.

Solution Consider the oil to be subdivided into disks of thickness Δy and radius x, as shown in Figure 7.52. Because the increment of force for each disk is given by its weight, you have

$$\begin{aligned}\Delta F &= \text{weight}\\ &= \left(\frac{50 \text{ pounds}}{\text{cubic foot}}\right)(\text{volume})\\ &= 50(\pi x^2 \Delta y) \ \text{ pounds.}\end{aligned}$$

For a circle of radius 8 and center at $(0, 8)$, you have

$$\begin{aligned}x^2 + (y-8)^2 &= 8^2\\ x^2 &= 16y - y^2\end{aligned}$$

and you can write the force increment as

$$\begin{aligned}\Delta F &= 50(\pi x^2 \Delta y)\\ &= 50\pi(16y - y^2)\,\Delta y.\end{aligned}$$

In Figure 7.52, note that a disk y feet from the bottom of the tank must be moved a distance of $(16 - y)$ feet. So, the increment of work is

$$\begin{aligned}\Delta W &= \Delta F(16-y)\\ &= 50\pi(16y - y^2)\,\Delta y(16-y)\\ &= 50\pi(256y - 32y^2 + y^3)\,\Delta y.\end{aligned}$$

Because the tank is half full, y ranges from 0 to 8, and the work required to empty the tank is

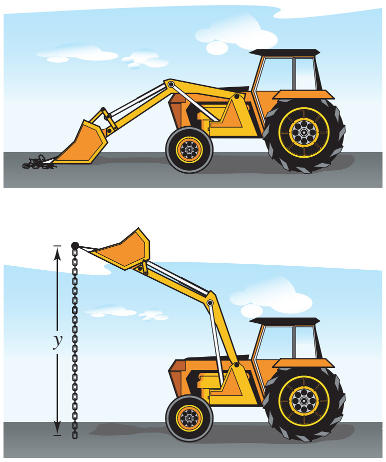
$$\begin{aligned}W &= \int_0^8 50\pi(256y - 32y^2 + y^3)\,dy\\ &= 50\pi\left[128y^2 - \frac{32}{3}y^3 + \frac{y^4}{4}\right]_0^8\\ &= 50\pi\left(\frac{11{,}264}{3}\right)\\ &\approx 589{,}782 \ \text{ foot-pounds.}\end{aligned}$$

Try It | Exploration A

To estimate the reasonableness of the result in Example 4, consider that the weight of the oil in the tank is

$$\begin{aligned}\left(\frac{1}{2}\right)(\text{volume})(\text{density}) &= \frac{1}{2}\left(\frac{4}{3}\pi 8^3\right)(50)\\ &\approx 53{,}616.5 \ \text{ pounds.}\end{aligned}$$

Lifting the entire half-tank of oil 8 feet would involve work of $8(53{,}616.5) \approx 428{,}932$ foot-pounds. Because the oil is actually lifted between 8 and 16 feet, it seems reasonable that the work done is 589,782 foot-pounds.



Work required to raise one end of the chain
Figure 7.53

EXAMPLE 5 Lifting a Chain

A 20-foot chain weighing 5 pounds per foot is lying coiled on the ground. How much work is required to raise one end of the chain to a height of 20 feet so that it is fully extended, as shown in Figure 7.53?

Solution Imagine that the chain is divided into small sections, each of length Δy. Then the weight of each section is the increment of force

$$\Delta F = (\text{weight}) = \left(\frac{5 \text{ pounds}}{\text{foot}}\right)(\text{length}) = 5\Delta y.$$

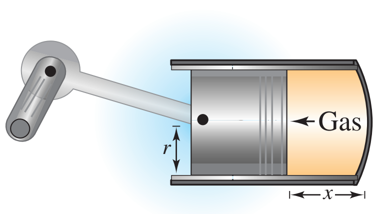
Because a typical section (initially on the ground) is raised to a height of y, the increment of work is

$$\Delta W = (\text{force increment})(\text{distance}) = (5\,\Delta y)y = 5y\,\Delta y.$$

Because y ranges from 0 to 20, the total work is

$$W = \int_0^{20} 5y\,dy = \frac{5y^2}{2}\bigg]_0^{20} = \frac{5(400)}{2} = 1000 \text{ foot-pounds.}$$

Try It Exploration A



Work done by expanding gas
Figure 7.54

In the next example you will consider a piston of radius r in a cylindrical casing, as shown in Figure 7.54. As the gas in the cylinder expands, the piston moves and work is done. If p represents the pressure of the gas (in pounds per square foot) against the piston head and V represents the volume of the gas (in cubic feet), the work increment involved in moving the piston Δx feet is

$$\Delta W = (\text{force})(\text{distance increment}) = F(\Delta x) = p(\pi r^2)\,\Delta x = p\,\Delta V.$$

So, as the volume of the gas expands from V_0 to V_1, the work done in moving the piston is

$$W = \int_{V_0}^{V_1} p\,dV.$$

Assuming the pressure of the gas to be inversely proportional to its volume, you have $p = k/V$ and the integral for work becomes

$$W = \int_{V_0}^{V_1} \frac{k}{V}\,dV.$$

EXAMPLE 6 Work Done by an Expanding Gas

A quantity of gas with an initial volume of 1 cubic foot and a pressure of 500 pounds per square foot expands to a volume of 2 cubic feet. Find the work done by the gas. (Assume that the pressure is inversely proportional to the volume.)

Solution Because $p = k/V$ and $p = 500$ when $V = 1$, you have $k = 500$. So, the work is

$$\begin{aligned} W &= \int_{V_0}^{V_1} \frac{k}{V}\,dV \\ &= \int_1^2 \frac{500}{V}\,dV \\ &= 500 \ln|V|\bigg]_1^2 \approx 346.6 \text{ foot-pounds.} \end{aligned}$$

Try It Exploration A

Section 7.6 Moments, Centers of Mass, and Centroids

- **Understand the definition of mass.**
- **Find the center of mass in a one-dimensional system.**
- **Find the center of mass in a two-dimensional system.**
- **Find the center of mass of a planar lamina.**
- **Use the Theorem of Pappus to find the volume of a solid of revolution.**

Mass

In this section you will study several important applications of integration that are related to **mass.** Mass is a measure of a body's resistance to changes in motion, and is independent of the particular gravitational system in which the body is located. However, because so many applications involving mass occur on Earth's surface, an object's mass is sometimes equated with its weight. This is not technically correct. Weight is a type of force and as such is dependent on gravity. Force and mass are related by the equation

$$\text{Force} = (\text{mass})(\text{acceleration}).$$

The table below lists some commonly used measures of mass and force, together with their conversion factors.

System of Measurement	Measure of Mass	Measure of Force
U.S.	Slug	Pound $=$ (slug)(ft/sec^2)
International	Kilogram	Newton $=$ (kilogram)(m/sec^2)
C-G-S	Gram	Dyne $=$ (gram)(cm/sec^2)

Conversions:

1 pound $=$ 4.448 newtons	1 slug $=$ 14.59 kilograms
1 newton $=$ 0.2248 pound	1 kilogram $=$ 0.06852 slug
1 dyne $=$ 0.000002248 pound	1 gram $=$ 0.00006852 slug
1 dyne $=$ 0.00001 newton	1 foot $=$ 0.3048 meter

EXAMPLE 1 Mass on the Surface of Earth

Find the mass (in slugs) of an object whose weight at sea level is 1 pound.

Solution Using 32 feet per second per second as the acceleration due to gravity produces

$$\begin{aligned}\text{Mass} &= \frac{\text{force}}{\text{acceleration}} &&\text{Force} = (\text{mass})(\text{acceleration})\\ &= \frac{1 \text{ pound}}{32 \text{ feet per second per second}}\\ &= 0.03125 \frac{\text{pound}}{\text{foot per second per second}}\\ &= 0.03125 \text{ slug.}\end{aligned}$$

Because many applications involving mass occur on Earth's surface, this amount of mass is called a **pound mass.**

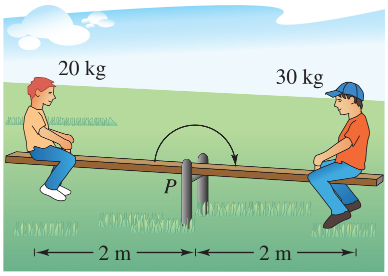
Try It Exploration A

Center of Mass in a One-Dimensional System

You will now consider two types of moments of a mass—the **moment about a point** and the **moment about a line.** To define these two moments, consider an idealized situation in which a mass m is concentrated at a point. If x is the distance between this point mass and another point P, the **moment of m about the point P is**

$$\text{Moment} = mx$$

and x is the **length of the moment arm.**



The seesaw will balance when the left and the right moments are equal.
Figure 7.55

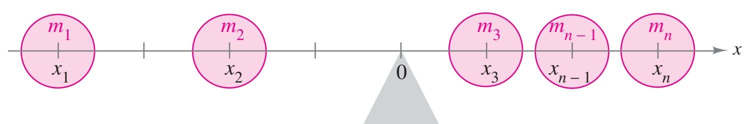
The concept of moment can be demonstrated simply by a seesaw, as shown in Figure 7.55. A child of mass 20 kilograms sits 2 meters to the left of fulcrum P, and an older child of mass 30 kilograms sits 2 meters to the right of P. From experience, you know that the seesaw will begin to rotate clockwise, moving the larger child down. This rotation occurs because the moment produced by the child on the left is less than the moment produced by the child on the right.

$$\begin{aligned} \text{Left moment} &= (20)(2) = 40 \text{ kilogram-meters} \\ \text{Right moment} &= (30)(2) = 60 \text{ kilogram-meters} \end{aligned}$$

To balance the seesaw, the two moments must be equal. For example, if the larger child moved to a position $\frac{4}{3}$ meters from the fulcrum, the seesaw would balance, because each child would produce a moment of 40 kilogram-meters.

To generalize this, you can introduce a coordinate line on which the origin corresponds to the fulcrum, as shown in Figure 7.56. Suppose several point masses are located on the x-axis. The measure of the tendency of this system to rotate about the origin is the **moment about the origin,** and it is defined as the sum of the n products $m_i x_i$.

$$M_0 = m_1x_1 + m_2x_2 + \cdot \cdot \cdot + m_nx_n$$



If $m_1x_1 + m_2x_2 + \cdot \cdot \cdot + m_nx_n = 0$, the system is in equilibrium.
Figure 7.56

If M_0 is 0, the system is said to be in **equilibrium.** The concept of equilibrium is demonstrated in the simulation below.

Simulation

For a system that is not in equilibrium, the **center of mass** is defined as the point $\bar{x}$ at which the fulcrum could be relocated to attain equilibrium. If the system were translated $\bar{x}$ units, each coordinate x_i would become $(x_i - \bar{x})$, and because the moment of the translated system is 0, you have

$$\sum_{i=1}^{n} m_i(x_i - \bar{x}) = \sum_{i=1}^{n} m_ix_i - \sum_{i=1}^{n} m_i\bar{x} = 0.$$

Solving for $\bar{x}$ produces

$$\bar{x} = \frac{\sum_{i=1}^{n} m_ix_i}{\sum_{i=1}^{n} m_i} = \frac{\text{moment of system about origin}}{\text{total mass of system}}.$$

If $m_1x_1 + m_2x_2 + \cdot \cdot \cdot + m_nx_n = 0$, the system is in equilibrium.

Moments and Center of Mass: One-Dimensional System

Let the point masses $m_1, m_2, \ldots, m_n$ be located at $x_1, x_2, \ldots, x_n$.

1. The **moment about the origin** is $M_0 = m_1x_1 + m_2x_2 + \cdots + m_nx_n$.
2. The **center of mass** is $\bar{x} = \dfrac{M_0}{m}$, where $m = m_1 + m_2 + \cdots + m_n$ is the **total mass** of the system.

EXAMPLE 2 The Center of Mass of a Linear System

Find the center of mass of the linear system shown in Figure 7.57.

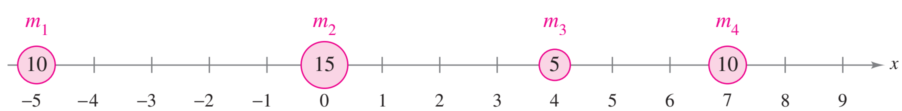


Figure 7.57

Solution The moment about the origin is

$$\begin{aligned} M_0 &= m_1x_1 + m_2x_2 + m_3x_3 + m_4x_4 \\ &= 10(-5) + 15(0) + 5(4) + 10(7) \\ &= -50 + 0 + 20 + 70 \\ &= 40. \end{aligned}$$

Because the total mass of the system is $m = 10 + 15 + 5 + 10 = 40$, the center of mass is

$$\bar{x} = \frac{M_0}{m} = \frac{40}{40} = 1.$$

Try It

Exploration A

NOTE In Example 2, where should you locate the fulcrum so that the point masses will be in equilibrium?

Rather than define the moment of a mass, you could define the moment of a *force*. In this context, the center of mass is called the **center of gravity.** Suppose that a system of point masses $m_1, m_2, \ldots, m_n$ is located at $x_1, x_2, \ldots, x_n$. Then, because force = (mass)(acceleration), the total force of the system is

$$\begin{aligned} F &= m_1a + m_2a + \cdots + m_na \\ &= ma. \end{aligned}$$

The **torque** (moment) about the origin is

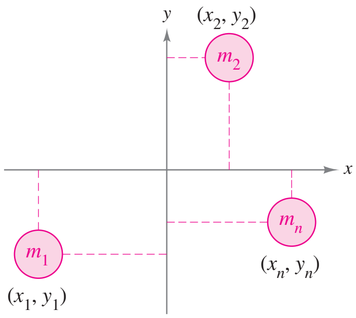
$$\begin{aligned} T_0 &= (m_1a)x_1 + (m_2a)x_2 + \cdots + (m_na)x_n \\ &= M_0a \end{aligned}$$

and the **center of gravity** is

$$\frac{T_0}{F} = \frac{M_0a}{ma} = \frac{M_0}{m} = \bar{x}.$$

So, the center of gravity and the center of mass have the same location.

Center of Mass in a Two-Dimensional System



In a two-dimensional system, there is a moment about the y-axis, M_y, and a moment about the x-axis, M_x.
Figure 7.58

You can extend the concept of moment to two dimensions by considering a system of masses located in the xy-plane at the points $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$, as shown in Figure 7.58. Rather than defining a single moment (with respect to the origin), two moments are defined—one with respect to the x-axis and one with respect to the y-axis.

Moments and Center of Mass: Two-Dimensional System

Let the point masses $m_1, m_2, \ldots, m_n$ be located at $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$.

1. The **moment about the y-axis** is $M_y = m_1x_1 + m_2x_2 + \cdots + m_nx_n$.
2. The **moment about the x-axis** is $M_x = m_1y_1 + m_2y_2 + \cdots + m_ny_n$.
3. The **center of mass** $(\bar{x}, \bar{y})$ (or **center of gravity**) is

$$\bar{x} = \frac{M_y}{m} \quad \text{and} \quad \bar{y} = \frac{M_x}{m}$$

where $m = m_1 + m_2 + \cdots + m_n$ is the **total mass** of the system.

The moment of a system of masses in the plane can be taken about any horizontal or vertical line. In general, the moment about a line is the sum of the product of the masses and the *directed distances* from the points to the line.

$$\text{Moment} = m_1(y_1 - b) + m_2(y_2 - b) + \cdots + m_n(y_n - b) \qquad \text{Horizontal line } y = b$$
$$\text{Moment} = m_1(x_1 - a) + m_2(x_2 - a) + \cdots + m_n(x_n - a) \qquad \text{Vertical line } x = a$$

EXAMPLE 3 The Center of Mass of a Two-Dimensional System

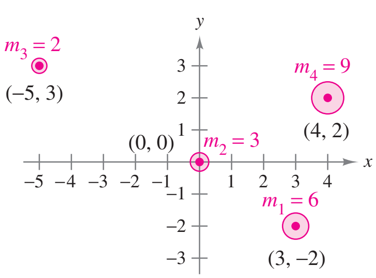


Figure 7.59

Find the center of mass of a system of point masses $m_1 = 6$, $m_2 = 3$, $m_3 = 2$, and $m_4 = 9$, located at

$$(3, -2), (0, 0), (-5, 3), \text{ and } (4, 2)$$

as shown in Figure 7.59.

Solution

$$\begin{aligned} m &= 6 + 3 + 2 + 9 = 20 && \text{Mass} \\ M_y &= 6(3) + 3(0) + 2(-5) + 9(4) = 44 && \text{Moment about } y\text{-axis} \\ M_x &= 6(-2) + 3(0) + 2(3) + 9(2) = 12 && \text{Moment about } x\text{-axis} \end{aligned}$$

So,

$$\bar{x} = \frac{M_y}{m} = \frac{44}{20} = \frac{11}{5}$$

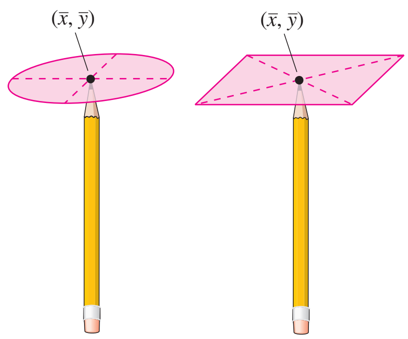
and

$$\bar{y} = \frac{M_x}{m} = \frac{12}{20} = \frac{3}{5}$$

and so the center of mass is $\left(\frac{11}{5}, \frac{3}{5}\right)$.

Try It

Exploration A



You can think of the center of mass $(\bar{x}, \bar{y})$ of a lamina as its balancing point. For a circular lamina, the center of mass is the center of the circle. For a rectangular lamina, the center of mass is the center of the rectangle.
Figure 7.60

Center of Mass of a Planar Lamina

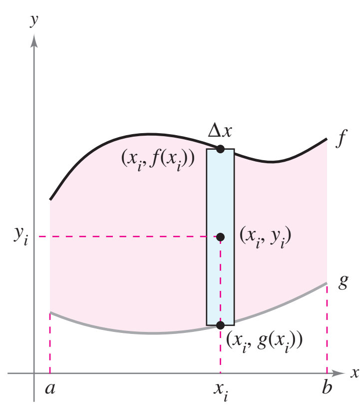
So far in this section you have assumed the total mass of a system to be distributed at discrete points in a plane or on a line. Now consider a thin, flat plate of material of constant density called a **planar lamina** (see Figure 7.60). **Density** is a measure of mass per unit of volume, such as grams per cubic centimeter. For planar laminas, however, density is considered to be a measure of mass per unit of area. Density is denoted by ρ, the lowercase Greek letter rho.

Consider an irregularly shaped planar lamina of uniform density ρ, bounded by the graphs of $y = f(x)$, $y = g(x)$, and $a \le x \le b$, as shown in Figure 7.61. The mass of this region is given by

$$\begin{aligned} m &= (\text{density})(\text{area}) \\ &= \rho \int_a^b [f(x) - g(x)]\,dx \\ &= \rho A \end{aligned}$$

where A is the area of the region. To find the center of mass of this lamina, partition the interval $[a, b]$ into n subintervals of equal width Δx. Let x_i be the center of the ith subinterval. You can approximate the portion of the lamina lying in the ith subinterval by a rectangle whose height is $h = f(x_i) - g(x_i)$. Because the density of the rectangle is ρ, its mass is

$$\begin{aligned} m_i &= (\text{density})(\text{area}) \\ &= \underbrace{\rho}_{\text{Density}}\,\underbrace{[f(x_i) - g(x_i)]}_{\text{Height}}\,\underbrace{\Delta x}_{\text{Width}}. \end{aligned}$$



Planar lamina of uniform density ρ
Figure 7.61

Now, considering this mass to be located at the center (x_i, y_i) of the rectangle, the directed distance from the x-axis to (x_i, y_i) is $y_i = [f(x_i) + g(x_i)]/2$. So, the moment of m_i about the x-axis is

$$\begin{aligned} \text{Moment} &= (\text{mass})(\text{distance}) \\ &= m_i y_i \\ &= \rho[f(x_i) - g(x_i)]\,\Delta x\left[\frac{f(x_i) + g(x_i)}{2}\right]. \end{aligned}$$

Summing the moments and taking the limit as $n \to \infty$ suggest the definitions below.

Moments and Center of Mass of a Planar Lamina

Let f and g be continuous functions such that $f(x) \ge g(x)$ on $[a, b]$, and consider the planar lamina of uniform density ρ bounded by the graphs of $y = f(x)$, $y = g(x)$, and $a \le x \le b$.

1. The **moments about the x- and y-axes** are

$$M_x = \rho \int_a^b \left[\frac{f(x) + g(x)}{2}\right][f(x) - g(x)]\,dx$$

$$M_y = \rho \int_a^b x[f(x) - g(x)]\,dx.$$

2. The **center of mass** $(\bar{x}, \bar{y})$ is given by $\bar{x} = \dfrac{M_y}{m}$ and $\bar{y} = \dfrac{M_x}{m}$, where $m = \rho \int_a^b [f(x) - g(x)]\,dx$ is the mass of the lamina.

EXAMPLE 4 The Center of Mass of a Planar Lamina

Find the center of mass of the lamina of uniform density ρ bounded by the graph of $f(x) = 4 - x^2$ and the x-axis.

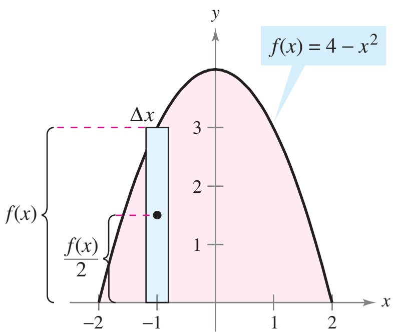


Figure 7.62

Editable Graph

Solution Because the center of mass lies on the axis of symmetry, you know that $\bar{x} = 0$. Moreover, the mass of the lamina is

$$\begin{aligned} m &= \rho \int_{-2}^{2} (4 - x^2)\,dx \\ &= \rho\left[4x - \frac{x^3}{3}\right]_{-2}^{2} \\ &= \frac{32\rho}{3}. \end{aligned}$$

To find the moment about the x-axis, place a representative rectangle in the region, as shown in Figure 7.62. The distance from the x-axis to the center of this rectangle is

$$y_i = \frac{f(x)}{2} = \frac{4 - x^2}{2}.$$

Because the mass of the representative rectangle is

$$\rho f(x)\,\Delta x = \rho(4 - x^2)\,\Delta x$$

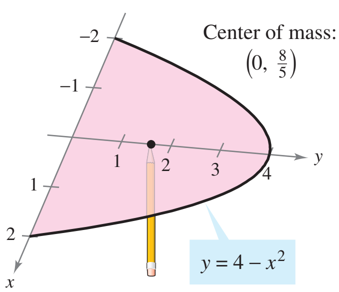
you have

$$\begin{aligned} M_x &= \rho \int_{-2}^{2} \frac{4 - x^2}{2}(4 - x^2)\,dx \\ &= \frac{\rho}{2} \int_{-2}^{2} (16 - 8x^2 + x^4)\,dx \\ &= \frac{\rho}{2}\left[16x - \frac{8x^3}{3} + \frac{x^5}{5}\right]_{-2}^{2} \\ &= \frac{256\rho}{15} \end{aligned}$$

and $\bar{y}$ is given by

$$\bar{y} = \frac{M_x}{m} = \frac{256\rho/15}{32\rho/3} = \frac{8}{5}.$$

So, the center of mass (the balancing point) of the lamina is $\left(0, \frac{8}{5}\right)$, as shown in Figure 7.63.



The center of mass is the balancing point.
Figure 7.63

Try It Exploration A Open Exploration

The density ρ in Example 4 is a common factor of both the moments and the mass, and as such divides out of the quotients representing the coordinates of the center of mass. So, the center of mass of a lamina of *uniform* density depends only on the shape of the lamina and not on its density. For this reason, the point

$(\bar{x}, \bar{y})$ Center of mass or centroid

is sometimes called the center of mass of a *region* in the plane, or the **centroid** of the region. In other words, to find the centroid of a region in the plane, you simply assume that the region has a constant density of $\rho = 1$ and compute the corresponding center of mass.

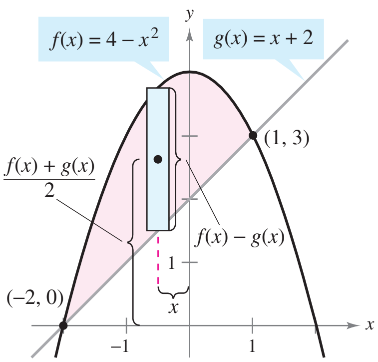


Figure 7.64

Editable Graph

EXPLORATION

Cut an irregular shape from a piece of cardboard.

a. Hold a pencil vertically and move the object on the pencil point until the centroid is located.

b. Divide the object into representative elements. Make the necessary measurements and numerically approximate the centroid. Compare your result with the result in part (a).

EXAMPLE 5 The Centroid of a Plane Region

Find the centroid of the region bounded by the graphs of $f(x) = 4 - x^2$ and $g(x) = x + 2$.

Solution The two graphs intersect at the points $(-2, 0)$ and $(1, 3)$, as shown in Figure 7.64. So, the area of the region is

$$A = \int_{-2}^{1} [f(x) - g(x)]\,dx = \int_{-2}^{1} (2 - x - x^2)\,dx = \frac{9}{2}.$$

The centroid $(\bar{x}, \bar{y})$ of the region has the following coordinates.

$$\bar{x} = \frac{1}{A}\int_{-2}^{1} x[(4 - x^2) - (x + 2)]\,dx = \frac{2}{9}\int_{-2}^{1} (-x^3 - x^2 + 2x)\,dx$$

$$= \frac{2}{9}\left[-\frac{x^4}{4} - \frac{x^3}{3} + x^2\right]_{-2}^{1} = -\frac{1}{2}$$

$$\bar{y} = \frac{1}{A}\int_{-2}^{1} \left[\frac{(4 - x^2) + (x + 2)}{2}\right][(4 - x^2) - (x + 2)]\,dx$$

$$= \frac{2}{9}\left(\frac{1}{2}\right)\int_{-2}^{1} (-x^2 + x + 6)(-x^2 - x + 2)\,dx$$

$$= \frac{1}{9}\int_{-2}^{1} (x^4 - 9x^2 - 4x + 12)\,dx$$

$$= \frac{1}{9}\left[\frac{x^5}{5} - 3x^3 - 2x^2 + 12x\right]_{-2}^{1} = \frac{12}{5}.$$

So, the centroid of the region is $(\bar{x}, \bar{y}) = \left(-\frac{1}{2}, \frac{12}{5}\right)$.

Try It Exploration A

For simple plane regions, you may be able to find the centroids without resorting to integration.

EXAMPLE 6 The Centroid of a Simple Plane Region

Find the centroid of the region shown in Figure 7.65(a).

Solution By superimposing a coordinate system on the region, as shown in Figure 7.65(b), you can locate the centroids of the three rectangles at

$$\left(\frac{1}{2}, \frac{3}{2}\right), \quad \left(\frac{5}{2}, \frac{1}{2}\right), \quad \text{and} \quad (5, 1).$$

Using these three points, you can find the centroid of the region.

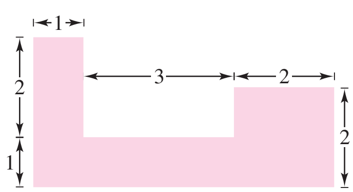
$$A = \text{area of region} = 3 + 3 + 4 = 10$$

$$\bar{x} = \frac{(1/2)(3) + (5/2)(3) + (5)(4)}{10} = \frac{29}{10} = 2.9$$

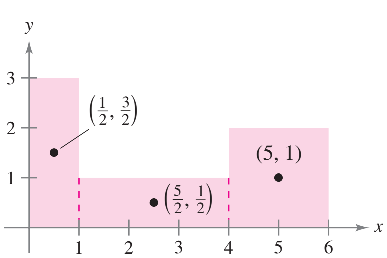
$$\bar{y} = \frac{(3/2)(3) + (1/2)(3) + (1)(4)}{10} = \frac{10}{10} = 1$$

So, the centroid of the region is $(2.9, 1)$.

Try It Exploration A



(a) Original region



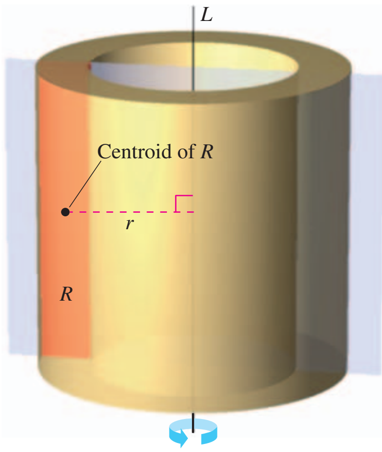
(b) The centroids of the three rectangles

Figure 7.65

NOTE In Example 6, notice that $(2.9, 1)$ is not the "average" of $\left(\frac{1}{2}, \frac{3}{2}\right)$, $\left(\frac{5}{2}, \frac{1}{2}\right)$, and $(5, 1)$.

Theorem of Pappus

The final topic in this section is a useful theorem credited to Pappus of Alexandria (ca. 300 A.D.), a Greek mathematician whose eight-volume *Mathematical Collection* is a record of much of classical Greek mathematics. The proof of this theorem is given in Section 14.4.



The volume V is $2\pi rA$, where A is the area of region R.
Figure 7.66

Rotatable Graph

THEOREM 7.1 The Theorem of Pappus

Let R be a region in a plane and let L be a line in the same plane such that L does not intersect the interior of R, as shown in Figure 7.66. If r is the distance between the centroid of R and the line, then the volume V of the solid of revolution formed by revolving R about the line is

$$V = 2\pi rA$$

where A is the area of R. (Note that $2\pi r$ is the distance traveled by the centroid as the region is revolved about the line.)

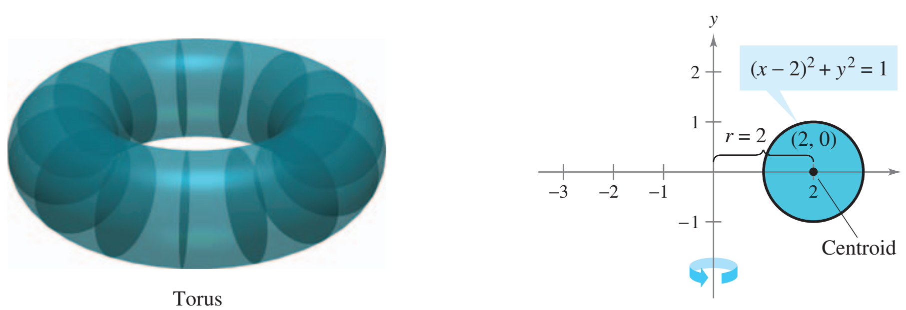
The Theorem of Pappus can be used to find the volume of a torus, as shown in the following example. Recall that a torus is a doughnut-shaped solid formed by revolving a circular region about a line that lies in the same plane as the circle (but does not intersect the circle).

EXAMPLE 7 Finding Volume by the Theorem of Pappus

Find the volume of the torus shown in Figure 7.67(a), which was formed by revolving the circular region bounded by

$$(x-2)^2 + y^2 = 1$$

about the y-axis, as shown in Figure 7.67(b).



(a) (b)
Figure 7.67

Rotatable Graph

Solution In Figure 7.67(b), you can see that the centroid of the circular region is $(2, 0)$. So, the distance between the centroid and the axis of revolution is $r = 2$. Because the area of the circular region is $A = \pi$, the volume of the torus is

$$\begin{aligned} V &= 2\pi rA \\ &= 2\pi(2)(\pi) \\ &= 4\pi^2 \\ &\approx 39.5. \end{aligned}$$

Try It Exploration A Exploration B

EXPLORATION

Use the shell method to show that the volume of the torus is given by

$$V = \int_1^3 4\pi x\sqrt{1-(x-2)^2}\,dx.$$

Evaluate this integral using a graphing utility. Does your answer agree with the one in Example 7?

Section 7.7

Fluid Pressure and Fluid Force

- **Find fluid pressure and fluid force.**

Fluid Pressure and Fluid Force

Swimmers know that the deeper an object is submerged in a fluid, the greater the pressure on the object. **Pressure** is defined as the force per unit of area over the surface of a body. For example, because a column of water that is 10 feet in height and 1 inch square weighs 4.3 pounds, the *fluid pressure* at a depth of 10 feet of water is 4.3 pounds per square inch.* At 20 feet, this would increase to 8.6 pounds per square inch, and in general the pressure is proportional to the depth of the object in the fluid.

Definition of Fluid Pressure

The **pressure** on an object at depth h in a liquid is

$$\text{Pressure} = P = wh$$

where w is the weight-density of the liquid per unit of volume.

Below are some common weight-densities of fluids in pounds per cubic foot.

Ethyl alcohol	49.4
Gasoline	41.0–43.0
Glycerin	78.6
Kerosene	51.2
Mercury	849.0
Seawater	64.0
Water	62.4

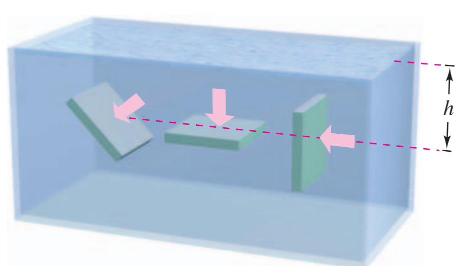
Blaise Pascal (1623–1662)

Pascal is well known for his work in many areas of mathematics and physics, and also for his influence on Leibniz. Although much of Pascal's work in calculus was intuitive and lacked the rigor of modern mathematics, he nevertheless anticipated many important results.

MathBio

When calculating fluid pressure, you can use an important (and rather surprising) physical law called **Pascal's Principle,** named after the French mathematician Blaise Pascal. Pascal's Principle states that the pressure exerted by a fluid at a depth h is transmitted equally *in all directions*. For example, in Figure 7.68, the pressure at the indicated depth is the same for all three objects. Because fluid pressure is given in terms of force per unit area ($P = F/A$), the fluid force on a *submerged horizontal* surface of area A is

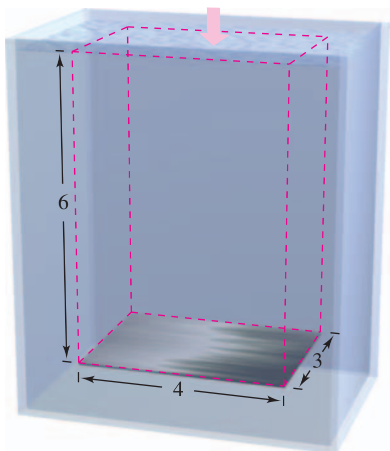
$$\text{Fluid force} = F = PA = (\text{pressure})(\text{area}).$$



The pressure at h is the same for all three objects.
Figure 7.68

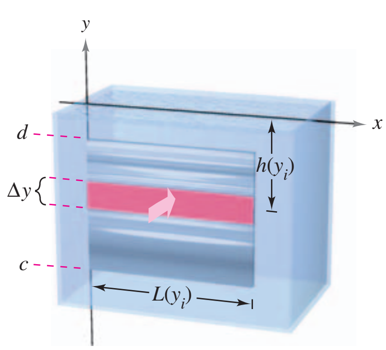
Rotatable Graph

* *The total pressure on an object in 10 feet of water would also include the pressure due to Earth's atmosphere. At sea level, atmospheric pressure is approximately 14.7 pounds per square inch.*



The fluid force on a horizontal metal sheet is equal to the fluid pressure times the area.
Figure 7.69

Rotatable Graph



Calculus methods must be used to find the fluid force on a vertical metal plate.
Figure 7.70

Rotatable Graph

EXAMPLE 1 Fluid Force on a Submerged Sheet

Find the fluid force on a rectangular metal sheet measuring 3 feet by 4 feet that is submerged in 6 feet of water, as shown in Figure 7.69.

Solution Because the weight-density of water is 62.4 pounds per cubic foot and the sheet is submerged in 6 feet of water, the fluid pressure is

$$\begin{aligned} P &= (62.4)(6) && P = wh \\ &= 374.4 \text{ pounds per square foot.} \end{aligned}$$

Because the total area of the sheet is $A = (3)(4) = 12$ square feet, the fluid force is

$$\begin{aligned} F = PA &= \left(374.4 \, \frac{\text{pounds}}{\text{square foot}}\right)(12 \text{ square feet}) \\ &= 4492.8 \text{ pounds.} \end{aligned}$$

This result is independent of the size of the body of water. The fluid force would be the same in a swimming pool or lake.

Try It   Exploration A

In Example 1, the fact that the sheet is rectangular and horizontal means that you do not need the methods of calculus to solve the problem. Consider a surface that is submerged vertically in a fluid. This problem is more difficult because the pressure is not constant over the surface.

Suppose a vertical plate is submerged in a fluid of weight-density w (per unit of volume), as shown in Figure 7.70. To determine the total force against *one side* of the region from depth c to depth d, you can subdivide the interval $[c, d]$ into n subintervals, each of width Δy. Next, consider the representative rectangle of width Δy and length $L(y_i)$, where y_i is in the ith subinterval. The force against this representative rectangle is

$$\begin{aligned} \Delta F_i &= w(\text{depth})(\text{area}) \\ &= wh(y_i)L(y_i)\,\Delta y. \end{aligned}$$

The force against n such rectangles is

$$\sum_{i=1}^{n} \Delta F_i = w \sum_{i=1}^{n} h(y_i)L(y_i)\,\Delta y.$$

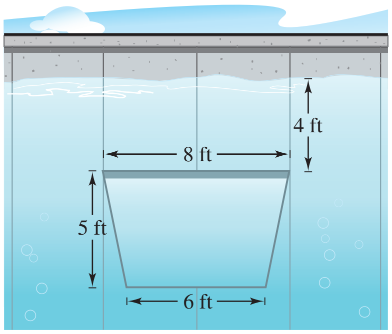
Note that w is considered to be constant and is factored out of the summation. Therefore, taking the limit as $\|\Delta\| \to 0$ $(n \to \infty)$ suggests the following definition.

Definition of Force Exerted by a Fluid

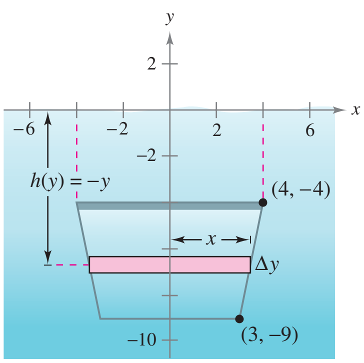
The **force F exerted by a fluid** of constant weight-density w (per unit of volume) against a submerged vertical plane region from $y = c$ to $y = d$ is

$$\begin{aligned} F &= w \lim_{\|\Delta\| \to 0} \sum_{i=1}^{n} h(y_i)L(y_i)\,\Delta y \\ &= w \int_c^d h(y)L(y)\,dy \end{aligned}$$

where $h(y)$ is the depth of the fluid at y and $L(y)$ is the horizontal length of the region at y.



(a) Water gate in a dam



(b) The fluid force against the gate
Figure 7.71

EXAMPLE 2 Fluid Force on a Vertical Surface

A vertical gate in a dam has the shape of an isosceles trapezoid 8 feet across the top and 6 feet across the bottom, with a height of 5 feet, as shown in Figure 7.71(a). What is the fluid force on the gate when the top of the gate is 4 feet below the surface of the water?

Solution In setting up a mathematical model for this problem, you are at liberty to locate the x- and y-axes in several different ways. A convenient approach is to let the y-axis bisect the gate and place the x-axis at the surface of the water, as shown in Figure 7.71(b). So, the depth of the water at y in feet is

$$\text{Depth} = h(y) = -y.$$

To find the length $L(y)$ of the region at y, find the equation of the line forming the right side of the gate. Because this line passes through the points $(3, -9)$ and $(4, -4)$, its equation is

$$\begin{aligned} y - (-9) &= \frac{-4 - (-9)}{4 - 3}(x - 3) \\ y + 9 &= 5(x - 3) \\ y &= 5x - 24 \\ x &= \frac{y + 24}{5}. \end{aligned}$$

In Figure 7.71(b) you can see that the length of the region at y is

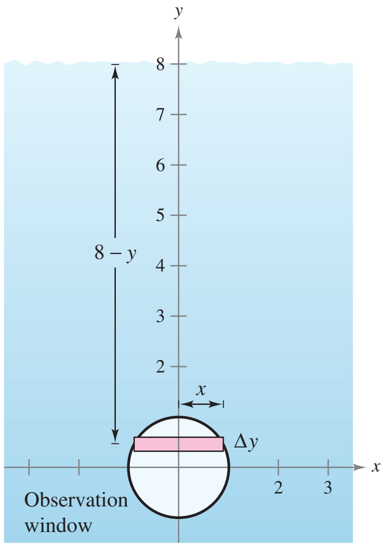
$$\begin{aligned} \text{Length} &= 2x \\ &= \frac{2}{5}(y + 24) \\ &= L(y). \end{aligned}$$

Finally, by integrating from $y = -9$ to $y = -4$, you can calculate the fluid force to be

$$\begin{aligned} F &= w\int_c^d h(y)L(y)\,dy \\ &= 62.4\int_{-9}^{-4}(-y)\left(\frac{2}{5}\right)(y + 24)\,dy \\ &= -62.4\left(\frac{2}{5}\right)\int_{-9}^{-4}(y^2 + 24y)\,dy \\ &= -62.4\left(\frac{2}{5}\right)\left[\frac{y^3}{3} + 12y^2\right]_{-9}^{-4} \\ &= -62.4\left(\frac{2}{5}\right)\left(\frac{-1675}{3}\right) \\ &= 13{,}936 \text{ pounds.} \end{aligned}$$

Try It | Exploration A | Open Exploration

NOTE In Example 2, the x-axis coincided with the surface of the water. This was convenient, but arbitrary. In choosing a coordinate system to represent a physical situation, you should consider various possibilities. Often you can simplify the calculations in a problem by locating the coordinate system to take advantage of special characteristics of the problem, such as symmetry.



The fluid force on the window
Figure 7.72

EXAMPLE 3 Fluid Force on a Vertical Surface

A circular observation window on a marine science ship has a radius of 1 foot, and the center of the window is 8 feet below water level, as shown in Figure 7.72. What is the fluid force on the window?

Solution To take advantage of symmetry, locate a coordinate system such that the origin coincides with the center of the window, as shown in Figure 7.72. The depth at y is then

$$\text{Depth} = h(y) = 8 - y.$$

The horizontal length of the window is $2x$, and you can use the equation for the circle, $x^2 + y^2 = 1$, to solve for x as follows.

$$\begin{aligned}\text{Length} &= 2x\\ &= 2\sqrt{1 - y^2} = L(y)\end{aligned}$$

Finally, because y ranges from -1 to 1, and using 64 pounds per cubic foot as the weight-density of seawater, you have

$$\begin{aligned}F &= w\int_c^d h(y)L(y)\,dy\\ &= 64\int_{-1}^{1}(8 - y)(2)\sqrt{1 - y^2}\,dy.\end{aligned}$$

Initially it looks as if this integral would be difficult to solve. However, if you break the integral into two parts and apply symmetry, the solution is simple.

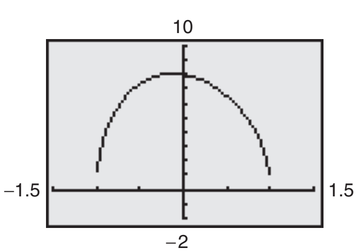
$$F = 64(16)\int_{-1}^{1}\sqrt{1 - y^2}\,dy - 64(2)\int_{-1}^{1} y\sqrt{1 - y^2}\,dy$$

The second integral is 0 (because the integrand is odd and the limits of integration are symmetric to the origin). Moreover, by recognizing that the first integral represents the area of a semicircle of radius 1, you obtain

$$\begin{aligned}F &= 64(16)\left(\frac{\pi}{2}\right) - 64(2)(0)\\ &= 512\pi\\ &\approx 1608.5 \text{ pounds.}\end{aligned}$$

So, the fluid force on the window is 1608.5 pounds.

Try It Exploration A



f is not differentiable at $x = \pm 1$.
Figure 7.73

TECHNOLOGY To confirm the result obtained in Example 3, you might have considered using Simpson's Rule to approximate the value of

$$128\int_{-1}^{1}(8 - x)\sqrt{1 - x^2}\,dx.$$

From the graph of

$$f(x) = (8 - x)\sqrt{1 - x^2}$$

however, you can see that f is not differentiable when $x = \pm 1$ (see Figure 7.73). This means that you cannot apply Theorem 4.19 from Section 4.6 to determine the potential error in Simpson's Rule. Without knowing the potential error, the approximation is of little value. Use a graphing utility to approximate the integral.

Section 8.1 Basic Integration Rules

- **Review procedures for fitting an integrand to one of the basic integration rules.**

Fitting Integrands to Basic Rules

In this chapter, you will study several integration techniques that greatly expand the set of integrals to which the basic integration rules can be applied. These rules are reviewed on page 520. A major step in solving any integration problem is recognizing which basic integration rule to use. As shown in Example 1, slight differences in the integrand can lead to very different solution techniques.

EXPLORATION

A Comparison of Three Similar Integrals Which, if any, of the following integrals can be evaluated using the 20 basic integration rules? For any that can be evaluated, do so. For any that can't, explain why.

a. $\displaystyle\int \frac{3}{\sqrt{1-x^2}}\,dx$

b. $\displaystyle\int \frac{3x}{\sqrt{1-x^2}}\,dx$

c. $\displaystyle\int \frac{3x^2}{\sqrt{1-x^2}}\,dx$

EXAMPLE 1 A Comparison of Three Similar Integrals

Find each integral.

a. $\displaystyle\int \frac{4}{x^2+9}\,dx$ **b.** $\displaystyle\int \frac{4x}{x^2+9}\,dx$ **c.** $\displaystyle\int \frac{4x^2}{x^2+9}\,dx$

Solution

a. Use the Arctangent Rule and let $u = x$ and $a = 3$.

$$\int \frac{4}{x^2+9}\,dx = 4\int \frac{1}{x^2+3^2}\,dx \qquad \text{Constant Multiple Rule}$$

$$= 4\left(\frac{1}{3}\arctan\frac{x}{3}\right) + C \qquad \text{Arctangent Rule}$$

$$= \frac{4}{3}\arctan\frac{x}{3} + C \qquad \text{Simplify.}$$

b. Here the Arctangent Rule does not apply because the numerator contains a factor of x. Consider the Log Rule and let $u = x^2 + 9$. Then $du = 2x\,dx$, and you have

$$\int \frac{4x}{x^2+9}\,dx = 2\int \frac{2x\,dx}{x^2+9} \qquad \text{Constant Multiple Rule}$$

$$= 2\int \frac{du}{u} \qquad \text{Substitution: } u = x^2+9$$

$$= 2\ln|u| + C = 2\ln(x^2+9) + C. \qquad \text{Log Rule}$$

c. Because the degree of the numerator is equal to the degree of the denominator, you should first use division to rewrite the improper rational function as the sum of a polynomial and a proper rational function.

$$\int \frac{4x^2}{x^2+9}\,dx = \int \left(4 - \frac{36}{x^2+9}\right) dx \qquad \text{Rewrite using long division.}$$

$$= \int 4\,dx - 36\int \frac{1}{x^2+9}\,dx \qquad \text{Write as two integrals.}$$

$$= 4x - 36\left(\frac{1}{3}\arctan\frac{x}{3}\right) + C \qquad \text{Integrate.}$$

$$= 4x - 12\arctan\frac{x}{3} + C \qquad \text{Simplify.}$$

NOTE Notice in Example 1(c) that some preliminary algebra is required before applying the rules for integration, and that subsequently more than one rule is needed to evaluate the resulting integral.

Try It | Exploration A | Exploration B | Open Exploration

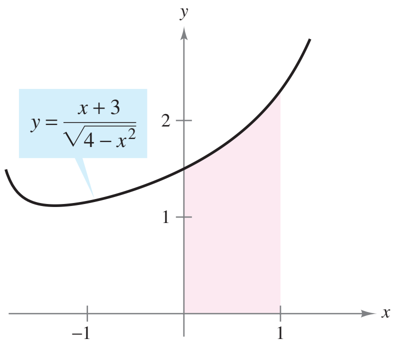
EXAMPLE 2 Using Two Basic Rules to Solve a Single Integral

Evaluate $\displaystyle\int_0^1 \frac{x+3}{\sqrt{4-x^2}}\,dx.$

Solution Begin by writing the integral as the sum of two integrals. Then apply the Power Rule and the Arcsine Rule as follows.

$$\begin{aligned}\int_0^1 \frac{x+3}{\sqrt{4-x^2}}\,dx &= \int_0^1 \frac{x}{\sqrt{4-x^2}}\,dx + \int_0^1 \frac{3}{\sqrt{4-x^2}}\,dx\\ &= -\frac{1}{2}\int_0^1 (4-x^2)^{-1/2}(-2x)\,dx + 3\int_0^1 \frac{1}{\sqrt{2^2-x^2}}\,dx\\ &= \left[-(4-x^2)^{1/2} + 3\arcsin\frac{x}{2}\right]_0^1\\ &= \left(-\sqrt{3}+\frac{\pi}{2}\right) - (-2+0)\\ &\approx 1.839\end{aligned}$$

See Figure 8.1.



The area of the region is approximately 1.839.
Figure 8.1

Editable Graph

Try It | Exploration A | Exploration B

TECHNOLOGY Simpson's Rule can be used to give a good approximation of the value of the integral in Example 2 (for $n = 10$, the approximation is 1.839). When using numerical integration, however, you should be aware that Simpson's Rule does not always give good approximations when one or both of the limits of integration are near a vertical asymptote. For instance, using the Fundamental Theorem of Calculus, you can obtain

$$\int_0^{1.99} \frac{x+3}{\sqrt{4-x^2}}\,dx \approx 6.213.$$

Applying Simpson's Rule (with $n = 10$) to this integral produces an approximation of 6.889.

EXAMPLE 3 A Substitution Involving $a^2 - u^2$

Find $\displaystyle\int \frac{x^2}{\sqrt{16-x^6}}\,dx.$

Solution Because the radical in the denominator can be written in the form

$$\sqrt{a^2-u^2} = \sqrt{4^2-(x^3)^2}$$

you can try the substitution $u = x^3$. Then $du = 3x^2\,dx$, and you have

$$\begin{aligned}\int \frac{x^2}{\sqrt{16-x^6}}\,dx &= \frac{1}{3}\int \frac{3x^2\,dx}{\sqrt{16-(x^3)^2}} && \text{Rewrite integral.}\\ &= \frac{1}{3}\int \frac{du}{\sqrt{4^2-u^2}} && \text{Substitution: } u = x^3\\ &= \frac{1}{3}\arcsin\frac{u}{4} + C && \text{Arcsine Rule}\\ &= \frac{1}{3}\arcsin\frac{x^3}{4} + C. && \text{Rewrite as a function of } x.\end{aligned}$$

Try It | Exploration A | Exploration B

STUDY TIP Rules 18, 19, and 20 of the basic integration rules on the next page all have expressions involving the sum or difference of two squares:

$a^2 - u^2$

$a^2 + u^2$

$u^2 - a^2$

With such an expression, consider the substitution $u = f(x)$, as in Example 3.

Surprisingly, two of the most commonly overlooked integration rules are the Log Rule and the Power Rule. Notice in the next two examples how these two integration rules can be disguised.

EXAMPLE 4 A Disguised Form of the Log Rule

Find $\displaystyle\int \frac{1}{1+e^x}\,dx$.

Solution The integral does not appear to fit any of the basic rules. However, the quotient form suggests the Log Rule. If you let $u = 1 + e^x$, then $du = e^x\,dx$. You can obtain the required du by adding and subtracting e^x in the numerator, as follows.

$$\int \frac{1}{1+e^x}\,dx = \int \frac{1+e^x-e^x}{1+e^x}\,dx \qquad \text{Add and subtract } e^x \text{ in numerator.}$$

$$= \int \left(\frac{1+e^x}{1+e^x} - \frac{e^x}{1+e^x}\right)dx \qquad \text{Rewrite as two fractions.}$$

$$= \int dx - \int \frac{e^x\,dx}{1+e^x} \qquad \text{Rewrite as two integrals.}$$

$$= x - \ln(1+e^x) + C \qquad \text{Integrate.}$$

Try It | Exploration A

NOTE There is usually more than one way to solve an integration problem. For instance, in Example 4, try integrating by multiplying the numerator and denominator by e^{-x} to obtain an integral of the form $-\int du/u$. See if you can get the same answer by this procedure. (Be careful: the answer will appear in a different form.)

EXAMPLE 5 A Disguised Form of the Power Rule

Find $\displaystyle\int (\cot x)[\ln(\sin x)]\,dx$.

Solution Again, the integral does not appear to fit any of the basic rules. However, considering the two primary choices for u [$u = \cot x$ and $u = \ln(\sin x)$], you can see that the second choice is the appropriate one because

$$u = \ln(\sin x) \qquad \text{and} \qquad du = \frac{\cos x}{\sin x}\,dx = \cot x\,dx.$$

So,

$$\int (\cot x)[\ln(\sin x)]\,dx = \int u\,du \qquad \text{Substitution: } u = \ln(\sin x)$$

$$= \frac{u^2}{2} + C \qquad \text{Integrate.}$$

$$= \frac{1}{2}[\ln(\sin x)]^2 + C. \qquad \text{Rewrite as a function of } x.$$

Try It | Exploration A | Technology

NOTE In Example 5, try *checking* that the derivative of

$$\frac{1}{2}[\ln(\sin x)]^2 + C$$

is the integrand of the original integral.

Review of Basic Integration Rules ($a > 0$)

1. $\int kf(u)\,du = k\int f(u)\,du$
2. $\int [f(u) \pm g(u)]\,du = \int f(u)\,du \pm \int g(u)\,du$
3. $\int du = u + C$
4. $\int u^n\,du = \dfrac{u^{n+1}}{n+1} + C,\ n \neq -1$
5. $\int \dfrac{du}{u} = \ln|u| + C$
6. $\int e^u\,du = e^u + C$
7. $\int a^u\,du = \left(\dfrac{1}{\ln a}\right)a^u + C$
8. $\int \sin u\,du = -\cos u + C$
9. $\int \cos u\,du = \sin u + C$
10. $\int \tan u\,du = -\ln|\cos u| + C$
11. $\int \cot u\,du = \ln|\sin u| + C$
12. $\int \sec u\,du = \ln|\sec u + \tan u| + C$
13. $\int \csc u\,du = -\ln|\csc u + \cot u| + C$
14. $\int \sec^2 u\,du = \tan u + C$
15. $\int \csc^2 u\,du = -\cot u + C$
16. $\int \sec u \tan u\,du = \sec u + C$
17. $\int \csc u \cot u\,du = -\csc u + C$
18. $\int \dfrac{du}{\sqrt{a^2-u^2}} = \arcsin\dfrac{u}{a} + C$
19. $\int \dfrac{du}{a^2+u^2} = \dfrac{1}{a}\arctan\dfrac{u}{a} + C$
20. $\int \dfrac{du}{u\sqrt{u^2-a^2}} = \dfrac{1}{a}\operatorname{arcsec}\dfrac{|u|}{a} + C$

Trigonometric identities can often be used to fit integrals to one of the basic integration rules.

EXAMPLE 6 Using Trigonometric Identities

Find $\int \tan^2 2x\, dx$.

TECHNOLOGY If you have access to a computer algebra system, try using it to evaluate the integrals in this section. Compare the *form* of the antiderivative given by the software with the form obtained by hand. Sometimes the forms will be the same, but often they will differ. For instance, why is the antiderivative $\ln 2x + C$ equivalent to the antiderivative $\ln x + C$?

Solution Note that $\tan^2 u$ is not in the list of basic integration rules. However, $\sec^2 u$ is in the list. This suggests the trigonometric identity $\tan^2 u = \sec^2 u - 1$. If you let $u = 2x$, then $du = 2\, dx$ and

$$\begin{aligned}
\int \tan^2 2x\, dx &= \frac{1}{2}\int \tan^2 u\, du && \text{Substitution: } u = 2x \\
&= \frac{1}{2}\int (\sec^2 u - 1)\, du && \text{Trigonometric identity} \\
&= \frac{1}{2}\int \sec^2 u\, du - \frac{1}{2}\int du && \text{Rewrite as two integrals.} \\
&= \frac{1}{2}\tan u - \frac{u}{2} + C && \text{Integrate.} \\
&= \frac{1}{2}\tan 2x - x + C. && \text{Rewrite as a function of } x.
\end{aligned}$$

Try It Exploration A

This section concludes with a summary of the common procedures for fitting integrands to the basic integration rules.

Procedures for Fitting Integrands to Basic Rules

Technique	*Example*
Expand (numerator).	$(1 + e^x)^2 = 1 + 2e^x + e^{2x}$
Separate numerator.	$\dfrac{1 + x}{x^2 + 1} = \dfrac{1}{x^2 + 1} + \dfrac{x}{x^2 + 1}$
Complete the square.	$\dfrac{1}{\sqrt{2x - x^2}} = \dfrac{1}{\sqrt{1 - (x - 1)^2}}$
Divide improper rational function.	$\dfrac{x^2}{x^2 + 1} = 1 - \dfrac{1}{x^2 + 1}$
Add and subtract terms in numerator.	$\dfrac{2x}{x^2 + 2x + 1} = \dfrac{2x + 2 - 2}{x^2 + 2x + 1} = \dfrac{2x + 2}{x^2 + 2x + 1} - \dfrac{2}{(x + 1)^2}$
Use trigonometric identities.	$\cot^2 x = \csc^2 x - 1$
Multiply and divide by Pythagorean conjugate.	$\dfrac{1}{1 + \sin x} = \left(\dfrac{1}{1 + \sin x}\right)\left(\dfrac{1 - \sin x}{1 - \sin x}\right) = \dfrac{1 - \sin x}{1 - \sin^2 x} = \dfrac{1 - \sin x}{\cos^2 x} = \sec^2 x - \dfrac{\sin x}{\cos^2 x}$

NOTE Remember that you can separate numerators but not denominators. Watch out for this common error when fitting integrands to basic rules.

$$\frac{1}{x^2 + 1} \neq \frac{1}{x^2} + \frac{1}{1} \qquad \text{Do not separate denominators.}$$

Section 8.2 Integration by Parts

- **Find an antiderivative using integration by parts.**
- **Use a tabular method to perform integration by parts.**

Integration by Parts

In this section you will study an important integration technique called **integration by parts.** This technique can be applied to a wide variety of functions and is particularly useful for integrands involving *products* of algebraic and transcendental functions. For instance, integration by parts works well with integrals such as

$$\int x \ln x \, dx, \quad \int x^2 e^x \, dx, \quad \text{and} \quad \int e^x \sin x \, dx.$$

Integration by parts is based on the formula for the derivative of a product

$$\frac{d}{dx}[uv] = u\frac{dv}{dx} + v\frac{du}{dx}$$
$$= uv' + vu'$$

where both u and v are differentiable functions of x. If u' and v' are continuous, you can integrate both sides of this equation to obtain

$$uv = \int uv' \, dx + \int vu' \, dx$$
$$= \int u \, dv + \int v \, du.$$

By rewriting this equation, you obtain the following theorem.

THEOREM 8.1 Integration by Parts

If u and v are functions of x and have continuous derivatives, then

$$\int u \, dv = uv - \int v \, du.$$

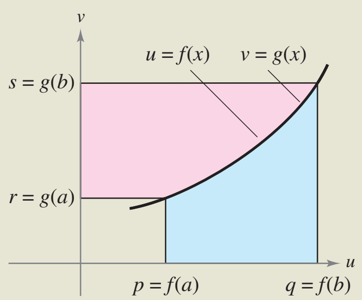
This formula expresses the original integral in terms of another integral. Depending on the choices of u and dv, it may be easier to evaluate the second integral than the original one. Because the choices of u and dv are critical in the integration by parts process, the following guidelines are provided.

Guidelines for Integration by Parts

1. Try letting dv be the most complicated portion of the integrand that fits a basic integration rule. Then u will be the remaining factor(s) of the integrand.
2. Try letting u be the portion of the integrand whose derivative is a function simpler than u. Then dv will be the remaining factor(s) of the integrand.

EXPLORATION

Proof Without Words Here is a different approach to proving the formula for integration by parts. Exercise taken from "Proof Without Words: Integration by Parts" by Roger B. Nelsen, *Mathematics Magazine*, April 1991, by permission of the author.



Area + Area $= qs - pr$

$$\int_r^s u \, dv + \int_q^p v \, du = \Big[uv\Big]_{(p,r)}^{(q,s)}$$

$$\int_r^s u \, dv = \Big[uv\Big]_{(p,r)}^{(q,s)} - \int_q^p v \, du$$

Explain how this graph proves the theorem. Which notation in this proof is unfamiliar? What do you think it means?

EXAMPLE 1 Integration by Parts

Find $\int xe^x\,dx$.

Solution To apply integration by parts, you need to write the integral in the form $\int u\,dv$. There are several ways to do this.

$$\int \underbrace{(x)}_{u}\underbrace{(e^x dx)}_{dv}, \quad \int \underbrace{(e^x)}_{u}\underbrace{(x\,dx)}_{dv}, \quad \int \underbrace{(1)}_{u}\underbrace{(xe^x\,dx)}_{dv}, \quad \int \underbrace{(xe^x)}_{u}\underbrace{(dx)}_{dv}$$

The guidelines on page 525 suggest choosing the first option because the derivative of $u = x$ is simpler than x, and $dv = e^x\,dx$ is the most complicated portion of the integrand that fits a basic integration formula.

$$dv = e^x\,dx \quad \Rightarrow \quad v = \int dv = \int e^x\,dx = e^x$$

$$u = x \quad \Rightarrow \quad du = dx$$

Now, integration by parts produces

$$\int u\,dv = uv - \int v\,du \qquad \text{Integration by parts formula}$$

$$\int xe^x\,dx = xe^x - \int e^x\,dx \qquad \text{Substitute.}$$

$$= xe^x - e^x + C. \qquad \text{Integrate.}$$

To check this, differentiate $xe^x - e^x + C$ to see that you obtain the original integrand.

NOTE In Example 1, note that it is not necessary to include a constant of integration when solving

$$v = \int e^x\,dx = e^x + C_1.$$

To illustrate this, replace $v = e^x$ by $v = e^x + C_1$ and apply integration by parts to see that you obtain the same result.

Try It **Exploration A** **Exploration B**

EXAMPLE 2 Integration by Parts

Find $\int x^2 \ln x\,dx$.

Solution In this case, x^2 is more easily integrated than $\ln x$. Furthermore, the derivative of $\ln x$ is simpler than $\ln x$. So, you should let $dv = x^2\,dx$.

$$dv = x^2\,dx \quad \Rightarrow \quad v = \int x^2\,dx = \frac{x^3}{3}$$

$$u = \ln x \quad \Rightarrow \quad du = \frac{1}{x}\,dx$$

Integration by parts produces

$$\int u\,dv = uv - \int v\,du \qquad \text{Integration by parts formula}$$

$$\int x^2 \ln x\,dx = \frac{x^3}{3}\ln x - \int \left(\frac{x^3}{3}\right)\left(\frac{1}{x}\right)dx \qquad \text{Substitute.}$$

$$= \frac{x^3}{3}\ln x - \frac{1}{3}\int x^2\,dx \qquad \text{Simplify.}$$

$$= \frac{x^3}{3}\ln x - \frac{x^3}{9} + C. \qquad \text{Integrate.}$$

You can check this result by differentiating.

$$\frac{d}{dx}\left[\frac{x^3}{3}\ln x - \frac{x^3}{9}\right] = \frac{x^3}{3}\left(\frac{1}{x}\right) + (\ln x)(x^2) - \frac{x^2}{3} = x^2 \ln x$$

TECHNOLOGY Try graphing

$$\int x^2 \ln x\,dx \quad \text{and} \quad \frac{x^3}{3}\ln x - \frac{x^3}{9}$$

on your graphing utility. Do you get the same graph? (This will take a while, so be patient.)

Try It **Exploration A** **Exploration B**

One surprising application of integration by parts involves integrands consisting of a single term, such as $\int \ln x\, dx$ or $\int \arcsin x\, dx$. In these cases, try letting $dv = dx$, as shown in the next example.

EXAMPLE 3 An Integrand with a Single Term

Evaluate $\displaystyle\int_0^1 \arcsin x\, dx$.

Solution Let $dv = dx$.

$$dv = dx \quad \Rightarrow \quad v = \int dx = x$$

$$u = \arcsin x \quad \Rightarrow \quad du = \frac{1}{\sqrt{1-x^2}}\, dx$$

Integration by parts now produces

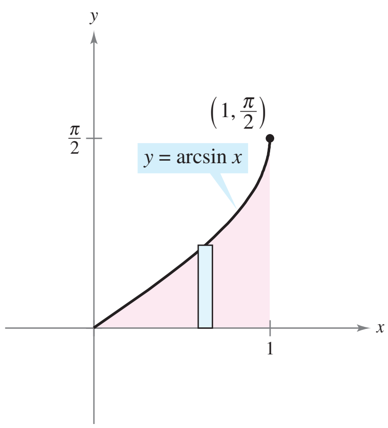
$$\int u\, dv = uv - \int v\, du \qquad \text{Integration by parts formula}$$

$$\begin{aligned} \int \arcsin x\, dx &= x \arcsin x - \int \frac{x}{\sqrt{1-x^2}}\, dx && \text{Substitute.} \\ &= x \arcsin x + \frac{1}{2}\int (1-x^2)^{-1/2}(-2x)\, dx && \text{Rewrite.} \\ &= x \arcsin x + \sqrt{1-x^2} + C. && \text{Integrate.} \end{aligned}$$

Using this antiderivative, you can evaluate the definite integral as follows.

$$\begin{aligned} \int_0^1 \arcsin x\, dx &= \Big[x \arcsin x + \sqrt{1-x^2}\Big]_0^1 \\ &= \frac{\pi}{2} - 1 \\ &\approx 0.571 \end{aligned}$$

The area represented by this definite integral is shown in Figure 8.2.



The area of the region is approximately 0.571.
Figure 8.2

Editable Graph

Try It   Exploration A   Exploration B

TECHNOLOGY Remember that there are two ways to use technology to evaluate a definite integral: (1) you can use a numerical approximation such as the Trapezoidal Rule or Simpson's Rule, or (2) you can use a computer algebra system to find the antiderivative and then apply the Fundamental Theorem of Calculus. Both methods have shortcomings. To find the possible error when using a numerical method, the integrand must have a second derivative (Trapezoidal Rule) or a fourth derivative (Simpson's Rule) in the interval of integration: the integrand in Example 3 fails to meet either of these requirements. To apply the Fundamental Theorem of Calculus, the symbolic integration utility must be able to find the antiderivative.

Which method would you use to evaluate

$$\int_0^1 \arctan x\, dx?$$

Which method would you use to evaluate

$$\int_0^1 \arctan x^2\, dx?$$

Some integrals require repeated use of the integration by parts formula.

EXAMPLE 4 Repeated Use of Integration by Parts

Find $\int x^2 \sin x\, dx$.

Solution The factors x^2 and $\sin x$ are equally easy to integrate. However, the derivative of x^2 becomes simpler, whereas the derivative of $\sin x$ does not. So, you should let $u = x^2$.

$$dv = \sin x\, dx \quad \Rightarrow \quad v = \int \sin x\, dx = -\cos x$$

$$u = x^2 \quad \Rightarrow \quad du = 2x\, dx$$

Now, integration by parts produces

$$\int x^2 \sin x\, dx = -x^2 \cos x + \int 2x \cos x\, dx.$$

First use of integration by parts

This first use of integration by parts has succeeded in simplifying the original integral, but the integral on the right still doesn't fit a basic integration rule. To evaluate that integral, you can apply integration by parts again. This time, let $u = 2x$.

$$dv = \cos x\, dx \quad \Rightarrow \quad v = \int \cos x\, dx = \sin x$$

$$u = 2x \quad \Rightarrow \quad du = 2\, dx$$

Now, integration by parts produces

$$\begin{aligned} \int 2x \cos x\, dx &= 2x \sin x - \int 2 \sin x\, dx \\ &= 2x \sin x + 2 \cos x + C. \end{aligned}$$

Second use of integration by parts

Combining these two results, you can write

$$\int x^2 \sin x\, dx = -x^2 \cos x + 2x \sin x + 2 \cos x + C.$$

Try It | Exploration A | Exploration B | Exploration C

When making repeated applications of integration by parts, you need to be careful not to interchange the substitutions in successive applications. For instance, in Example 4, the first substitution was $u = x^2$ and $dv = \sin x\, dx$. If, in the second application, you had switched the substitution to $u = \cos x$ and $dv = 2x$, you would have obtained

$$\begin{aligned} \int x^2 \sin x\, dx &= -x^2 \cos x + \int 2x \cos x\, dx \\ &= -x^2 \cos x + x^2 \cos x + \int x^2 \sin x\, dx \\ &= \int x^2 \sin x\, dx \end{aligned}$$

thereby undoing the previous integration and returning to the *original* integral. When making repeated applications of integration by parts, you should also watch for the appearance of a *constant multiple* of the original integral. For instance, this occurs when you use integration by parts to evaluate $\int e^x \cos 2x\, dx$, and also occurs in the next example.

EXPLORATION

Try to find

$$\int e^x \cos 2x\, dx$$

by letting $u = \cos 2x$ and $dv = e^x\, dx$ in the first substitution. For the second substitution, let $u = \sin 2x$ and $dv = e^x\, dx$.

NOTE The integral in Example 5 is an important one. In Section 8.4 (Example 5), you will see that it is used to find the arc length of a parabolic segment.

STUDY TIP The trigonometric identities

$$\sin^2 x = \frac{1 - \cos 2x}{2}$$

$$\cos^2 x = \frac{1 + \cos 2x}{2}$$

play an important role in this chapter.

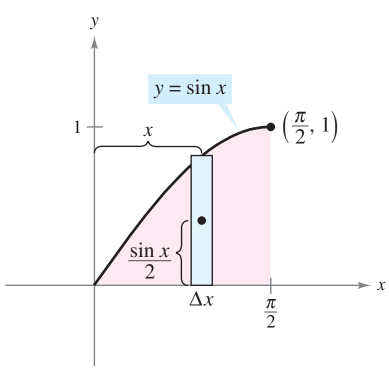


Figure 8.3

Editable Graph

EXAMPLE 5 Integration by Parts

Find $\int \sec^3 x\, dx$.

Solution The most complicated portion of the integrand that can be easily integrated is $\sec^2 x$, so you should let $dv = \sec^2 x\, dx$ and $u = \sec x$.

$$dv = \sec^2 x\, dx \quad \Rightarrow \quad v = \int \sec^2 x\, dx = \tan x$$

$$u = \sec x \quad \Rightarrow \quad du = \sec x \tan x\, dx$$

Integration by parts produces

$$\int u\, dv = uv - \int v\, du \qquad \text{Integration by parts formula}$$

$$\int \sec^3 x\, dx = \sec x \tan x - \int \sec x \tan^2 x\, dx \qquad \text{Substitute.}$$

$$\int \sec^3 x\, dx = \sec x \tan x - \int \sec x(\sec^2 x - 1)\, dx \qquad \text{Trigonometric identity}$$

$$\int \sec^3 x\, dx = \sec x \tan x - \int \sec^3 x\, dx + \int \sec x\, dx \qquad \text{Rewrite.}$$

$$2\int \sec^3 x\, dx = \sec x \tan x + \int \sec x\, dx \qquad \text{Collect like integrals.}$$

$$\int \sec^3 x\, dx = \frac{1}{2}\sec x \tan x + \frac{1}{2}\ln|\sec x + \tan x| + C. \qquad \text{Integrate and divide by 2.}$$

Try It | Exploration A | Exploration B

EXAMPLE 6 Finding a Centroid

A machine part is modeled by the region bounded by the graph of $y = \sin x$ and the x-axis, $0 \le x \le \pi/2$, as shown in Figure 8.3. Find the centroid of this region.

Solution Begin by finding the area of the region.

$$A = \int_0^{\pi/2} \sin x\, dx = \Big[-\cos x\Big]_0^{\pi/2} = 1$$

Now, you can find the coordinates of the centroid as follows.

$$\bar{y} = \frac{1}{A}\int_0^{\pi/2} \frac{\sin x}{2}(\sin x)\, dx = \frac{1}{4}\int_0^{\pi/2} (1 - \cos 2x)\, dx = \frac{1}{4}\left[x - \frac{\sin 2x}{2}\right]_0^{\pi/2} = \frac{\pi}{8}$$

You can evaluate the integral for $\bar{x}$, $(1/A)\int_0^{\pi/2} x \sin x\, dx$, with integration by parts. To do this, let $dv = \sin x\, dx$ and $u = x$. This produces $v = -\cos x$ and $du = dx$, and you can write

$$\int x \sin x\, dx = -x \cos x + \int \cos x\, dx$$

$$= -x \cos x + \sin x + C.$$

Finally, you can determine $\bar{x}$ to be

$$\bar{x} = \frac{1}{A}\int_0^{\pi/2} x \sin x\, dx = \Big[-x \cos x + \sin x\Big]_0^{\pi/2} = 1.$$

So, the centroid of the region is $(1, \pi/8)$.

Try It | Exploration A

As you gain experience in using integration by parts, your skill in determining u and dv will increase. The following summary lists several common integrals with suggestions for the choices of u and dv.

STUDY TIP You can use the acronym LIATE as a guideline for choosing u in integration by parts. In order, check the integrand for the following.

Is there a Logarithmic part?

Is there an Inverse trigonometric part?

Is there an Algebraic part?

Is there a Trigonometric part?

Is there an Exponential part?

Summary of Common Integrals Using Integration by Parts

1. For integrals of the form

$$\int x^n e^{ax}\,dx, \qquad \int x^n \sin ax\,dx, \qquad \text{or} \qquad \int x^n \cos ax\,dx$$

let $u = x^n$ and let $dv = e^{ax}\,dx$, $\sin ax\,dx$, or $\cos ax\,dx$.

2. For integrals of the form

$$\int x^n \ln x\,dx, \qquad \int x^n \arcsin ax\,dx, \qquad \text{or} \qquad \int x^n \arctan ax\,dx$$

let $u = \ln x$, $\arcsin ax$, or $\arctan ax$ and let $dv = x^n\,dx$.

3. For integrals of the form

$$\int e^{ax} \sin bx\,dx \qquad \text{or} \qquad \int e^{ax} \cos bx\,dx$$

let $u = \sin bx$ or $\cos bx$ and let $dv = e^{ax}\,dx$.

Tabular Method

In problems involving repeated applications of integration by parts, a tabular method, illustrated in Example 7, can help to organize the work. This method works well for integrals of the form $\int x^n \sin ax\,dx$, $\int x^n \cos ax\,dx$, and $\int x^n e^{ax}\,dx$.

EXAMPLE 7 Using the Tabular Method

Find $\displaystyle\int x^2 \sin 4x\,dx$.

Solution Begin as usual by letting $u = x^2$ and $dv = v'\,dx = \sin 4x\,dx$. Next, create a table consisting of three columns, as shown.

Alternate Signs	*u and Its Derivatives*	*v′ and Its Antiderivatives*
$+$	x^2	$\sin 4x$
$-$	$2x$	$-\frac{1}{4}\cos 4x$
$+$	2	$-\frac{1}{16}\sin 4x$
$-$	0	$\frac{1}{64}\cos 4x$

Differentiate until you obtain 0 as a derivative.

FOR FURTHER INFORMATION For more information on the tabular method, see the article "Tabular Integration by Parts" by David Horowitz in *The College Mathematics Journal*, and the article "More on Tabular Integration by Parts" by Leonard Gillman in *The College Mathematics Journal*.

MathArticle

MathArticle

The solution is obtained by adding the signed products of the diagonal entries:

$$\int x^2 \sin 4x\,dx = -\frac{1}{4}x^2 \cos 4x + \frac{1}{8}x \sin 4x + \frac{1}{32}\cos 4x + C.$$

Try It

Open Exploration

Section 8.3 Trigonometric Integrals

- **Solve trigonometric integrals involving powers of sine and cosine.**
- **Solve trigonometric integrals involving powers of secant and tangent.**
- **Solve trigonometric integrals involving sine-cosine products with different angles.**

Integrals Involving Powers of Sine and Cosine

SHEILA SCOTT MACINTYRE (1910–1960)

Sheila Scott Macintyre published her first paper on the asymptotic periods of integral functions in 1935. She completed her doctorate work at Aberdeen University, where she taught. In 1958 she accepted a visiting research fellowship at the University of Cincinnati.

In this section you will study techniques for evaluating integrals of the form

$$\int \sin^m x \cos^n x \, dx \qquad \text{and} \qquad \int \sec^m x \tan^n x \, dx$$

where either m or n is a positive integer. To find antiderivatives for these forms, try to break them into combinations of trigonometric integrals to which you can apply the Power Rule.

For instance, you can evaluate $\int \sin^5 x \cos x \, dx$ with the Power Rule by letting $u = \sin x$. Then, $du = \cos x \, dx$ and you have

$$\int \sin^5 x \cos x \, dx = \int u^5 \, du = \frac{u^6}{6} + C = \frac{\sin^6 x}{6} + C.$$

To break up $\int \sin^m x \cos^n x \, dx$ into forms to which you can apply the Power Rule, use the following identities.

$$\sin^2 x + \cos^2 x = 1 \qquad \text{Pythagorean identity}$$

$$\sin^2 x = \frac{1 - \cos 2x}{2} \qquad \text{Half-angle identity for } \sin^2 x$$

$$\cos^2 x = \frac{1 + \cos 2x}{2} \qquad \text{Half-angle identity for } \cos^2 x$$

Guidelines for Evaluating Integrals Involving Sine and Cosine

1. If the power of the sine is odd and positive, save one sine factor and convert the remaining factors to cosines. Then, expand and integrate.

$$\int \overbrace{\sin^{2k+1}}^{\text{Odd}} x \cos^n x \, dx = \int \overbrace{(\sin^2 x)^k}^{\text{Convert to cosines}} \cos^n x \overbrace{\sin x \, dx}^{\text{Save for } du} = \int (1 - \cos^2 x)^k \cos^n x \sin x \, dx$$

2. If the power of the cosine is odd and positive, save one cosine factor and convert the remaining factors to sines. Then, expand and integrate.

$$\int \sin^m x \overbrace{\cos^{2k+1}}^{\text{Odd}} x \, dx = \int \sin^m x \overbrace{(\cos^2 x)^k}^{\text{Convert to sines}} \overbrace{\cos x \, dx}^{\text{Save for } du} = \int \sin^m x \, (1 - \sin^2 x)^k \cos x \, dx$$

3. If the powers of both the sine and cosine are even and nonnegative, make repeated use of the identities

$$\sin^2 x = \frac{1 - \cos 2x}{2} \qquad \text{and} \qquad \cos^2 x = \frac{1 + \cos 2x}{2}$$

to convert the integrand to odd powers of the cosine. Then proceed as in guideline 2.

TECHNOLOGY Use a computer algebra system to find the integral in Example 1. You should obtain

$$\int \sin^3 x \cos^4 x\,dx =$$

$$-\cos^5 x\left(\frac{1}{7}\sin^2 x + \frac{2}{35}\right) + C.$$

Is this equivalent to the result obtained in Example 1?

EXAMPLE 1 Power of Sine Is Odd and Positive

Find $\int \sin^3 x \cos^4 x\,dx$.

Solution Because you expect to use the Power Rule with $u = \cos x$, *save one sine factor* to form du and convert the remaining sine factors to cosines.

$$\begin{aligned}
\int \sin^3 x \cos^4 x\,dx &= \int \sin^2 x \cos^4 x(\sin x)\,dx && \text{Rewrite.}\\
&= \int (1-\cos^2 x)\cos^4 x \sin x\,dx && \text{Trigonometric identity}\\
&= \int (\cos^4 x - \cos^6 x)\sin x\,dx && \text{Multiply.}\\
&= \int \cos^4 x \sin x\,dx - \int \cos^6 x \sin x\,dx && \text{Rewrite.}\\
&= -\int \cos^4 x(-\sin x)\,dx + \int \cos^6 x(-\sin x)\,dx\\
&= -\frac{\cos^5 x}{5} + \frac{\cos^7 x}{7} + C && \text{Integrate.}
\end{aligned}$$

Try It | Exploration A | Exploration B

Exploration C | Technology

In Example 1, *both* of the powers m and n happened to be positive integers. However, the same strategy will work as long as either m or n is odd and positive. For instance, in the next example the power of the cosine is 3, but the power of the sine is $-\frac{1}{2}$.

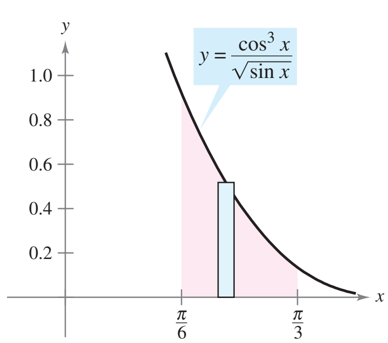
EXAMPLE 2 Power of Cosine Is Odd and Positive

Evaluate $\displaystyle\int_{\pi/6}^{\pi/3} \frac{\cos^3 x}{\sqrt{\sin x}}\,dx$.

Solution Because you expect to use the Power Rule with $u = \sin x$, *save one cosine factor* to form du and convert the remaining cosine factors to sines.

$$\begin{aligned}
\int_{\pi/6}^{\pi/3} \frac{\cos^3 x}{\sqrt{\sin x}}\,dx &= \int_{\pi/6}^{\pi/3} \frac{\cos^2 x \cos x}{\sqrt{\sin x}}\,dx\\
&= \int_{\pi/6}^{\pi/3} \frac{(1-\sin^2 x)(\cos x)}{\sqrt{\sin x}}\,dx\\
&= \int_{\pi/6}^{\pi/3} \left[(\sin x)^{-1/2}\cos x - (\sin x)^{3/2}\cos x\right]dx\\
&= \left[\frac{(\sin x)^{1/2}}{1/2} - \frac{(\sin x)^{5/2}}{5/2}\right]_{\pi/6}^{\pi/3}\\
&= 2\left(\frac{\sqrt{3}}{2}\right)^{1/2} - \frac{2}{5}\left(\frac{\sqrt{3}}{2}\right)^{5/2} - \sqrt{2} + \frac{\sqrt{32}}{80}\\
&\approx 0.239
\end{aligned}$$

Figure 8.4 shows the region whose area is represented by this integral.



The area of the region is approximately 0.239.
Figure 8.4

Editable Graph

Try It | Exploration A | Open Exploration

EXAMPLE 3 Power of Cosine Is Even and Nonnegative

Find $\int \cos^4 x\, dx$.

Solution Because m and n are both even and nonnegative $(m = 0)$, you can replace $\cos^4 x$ by $[(1 + \cos 2x)/2]^2$.

$$\begin{aligned}
\int \cos^4 x\, dx &= \int \left(\frac{1 + \cos 2x}{2}\right)^2 dx \\
&= \int \left(\frac{1}{4} + \frac{\cos 2x}{2} + \frac{\cos^2 2x}{4}\right) dx \\
&= \int \left[\frac{1}{4} + \frac{\cos 2x}{2} + \frac{1}{4}\left(\frac{1 + \cos 4x}{2}\right)\right] dx \\
&= \frac{3}{8}\int dx + \frac{1}{4}\int 2\cos 2x\, dx + \frac{1}{32}\int 4\cos 4x\, dx \\
&= \frac{3x}{8} + \frac{\sin 2x}{4} + \frac{\sin 4x}{32} + C
\end{aligned}$$

Use a symbolic differentiation utility to verify this. Can you simplify the derivative to obtain the original integrand?

Try It **Exploration A** **Exploration B**

In Example 3, if you were to evaluate the definite integral from 0 to $\pi/2$, you would obtain

$$\begin{aligned}
\int_0^{\pi/2} \cos^4 x\, dx &= \left[\frac{3x}{8} + \frac{\sin 2x}{4} + \frac{\sin 4x}{32}\right]_0^{\pi/2} \\
&= \left(\frac{3\pi}{16} + 0 + 0\right) - (0 + 0 + 0) \\
&= \frac{3\pi}{16}.
\end{aligned}$$

Note that the only term that contributes to the solution is $3x/8$. This observation is generalized in the following formulas developed by John Wallis.

Wallis's Formulas

1. If n is odd $(n \geq 3)$, then

$$\int_0^{\pi/2} \cos^n x\, dx = \left(\frac{2}{3}\right)\left(\frac{4}{5}\right)\left(\frac{6}{7}\right) \cdots \left(\frac{n-1}{n}\right).$$

2. If n is even $(n \geq 2)$, then

$$\int_0^{\pi/2} \cos^n x\, dx = \left(\frac{1}{2}\right)\left(\frac{3}{4}\right)\left(\frac{5}{6}\right) \cdots \left(\frac{n-1}{n}\right)\left(\frac{\pi}{2}\right).$$

These formulas are also valid if $\cos^n x$ is replaced by $\sin^n x$. (You are asked to prove both formulas in Exercise 104.)

John Wallis (1616–1703)

Wallis did much of his work in calculus prior to Newton and Leibniz, and he influenced the thinking of both of these men. Wallis is also credited with introducing the present symbol (∞) for infinity.

MathBio

Integrals Involving Powers of Secant and Tangent

The following guidelines can help you evaluate integrals of the form

$$\int \sec^m x \tan^n x \, dx.$$

Guidelines for Evaluating Integrals Involving Secant and Tangent

1. If the power of the secant is even and positive, save a secant-squared factor and convert the remaining factors to tangents. Then expand and integrate.

$$\int \sec^{\overbrace{2k}^{\text{Even}}} x \tan^n x \, dx = \int \overbrace{(\sec^2 x)^{k-1}}^{\text{Convert to tangents}} \tan^n x \overbrace{\sec^2 x \, dx}^{\text{Save for } du} = \int (1 + \tan^2 x)^{k-1} \tan^n x \sec^2 x \, dx$$

2. If the power of the tangent is odd and positive, save a secant-tangent factor and convert the remaining factors to secants. Then expand and integrate.

$$\int \sec^m x \tan^{\overbrace{2k+1}^{\text{Odd}}} x \, dx = \int \sec^{m-1} x \overbrace{(\tan^2 x)^k}^{\text{Convert to secants}} \overbrace{\sec x \tan x \, dx}^{\text{Save for } du} = \int \sec^{m-1} x (\sec^2 x - 1)^k \sec x \tan x \, dx$$

3. If there are no secant factors and the power of the tangent is even and positive, convert a tangent-squared factor to a secant-squared factor, then expand and repeat if necessary.

$$\int \tan^n x \, dx = \int \tan^{n-2} x \overbrace{(\tan^2 x)}^{\text{Convert to secants}} \, dx = \int \tan^{n-2} x (\sec^2 x - 1) \, dx$$

4. If the integral is of the form $\int \sec^m x \, dx$, where m is odd and positive, use integration by parts, as illustrated in Example 5 in the preceding section.

5. If none of the first four guidelines applies, try converting to sines and cosines.

EXAMPLE 4 Power of Tangent Is Odd and Positive

Find $\displaystyle\int \frac{\tan^3 x}{\sqrt{\sec x}} \, dx.$

Solution Because you expect to use the Power Rule with $u = \sec x$, *save a factor of* ($\sec x \tan x$) to form du and convert the remaining tangent factors to secants.

$$\begin{aligned}
\int \frac{\tan^3 x}{\sqrt{\sec x}} \, dx &= \int (\sec x)^{-1/2} \tan^3 x \, dx \\
&= \int (\sec x)^{-3/2} (\tan^2 x)(\sec x \tan x) \, dx \\
&= \int (\sec x)^{-3/2} (\sec^2 x - 1)(\sec x \tan x) \, dx \\
&= \int [(\sec x)^{1/2} - (\sec x)^{-3/2}](\sec x \tan x) \, dx \\
&= \frac{2}{3}(\sec x)^{3/2} + 2(\sec x)^{-1/2} + C
\end{aligned}$$

Try It Exploration A Exploration B

NOTE In Example 5, the power of the tangent is odd and positive. So, you could also find the integral using the procedure described in guideline 2 on page 537. In Exercise 85, you are asked to show that the results obtained by these two procedures differ only by a constant.

EXAMPLE 5 Power of Secant Is Even and Positive

Find $\int \sec^4 3x \tan^3 3x \, dx$.

Solution Let $u = \tan 3x$, then $du = 3 \sec^2 3x \, dx$ and you can write

$$\begin{aligned}
\int \sec^4 3x \tan^3 3x \, dx &= \int \sec^2 3x \tan^3 3x(\sec^2 3x) \, dx \\
&= \int (1 + \tan^2 3x) \tan^3 3x(\sec^2 3x) \, dx \\
&= \frac{1}{3}\int (\tan^3 3x + \tan^5 3x)(3 \sec^2 3x) \, dx \\
&= \frac{1}{3}\left(\frac{\tan^4 3x}{4} + \frac{\tan^6 3x}{6}\right) + C \\
&= \frac{\tan^4 3x}{12} + \frac{\tan^6 3x}{18} + C.
\end{aligned}$$

Try It | Exploration A

EXAMPLE 6 Power of Tangent Is Even

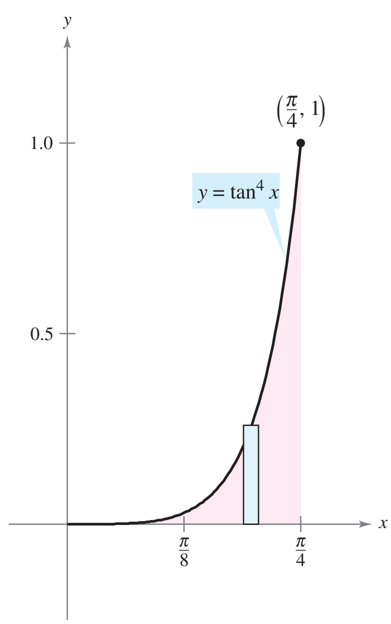
Evaluate $\int_0^{\pi/4} \tan^4 x \, dx$.

Solution Because there are no secant factors, you can begin by converting a tangent-squared factor to a secant-squared factor.

$$\begin{aligned}
\int \tan^4 x \, dx &= \int \tan^2 x(\tan^2 x) \, dx \\
&= \int \tan^2 x(\sec^2 x - 1) \, dx \\
&= \int \tan^2 x \sec^2 x \, dx - \int \tan^2 x \, dx \\
&= \int \tan^2 x \sec^2 x \, dx - \int (\sec^2 x - 1) \, dx \\
&= \frac{\tan^3 x}{3} - \tan x + x + C
\end{aligned}$$

You can evaluate the definite integral as follows.

$$\begin{aligned}
\int_0^{\pi/4} \tan^4 x \, dx &= \left[\frac{\tan^3 x}{3} - \tan x + x\right]_0^{\pi/4} \\
&= \frac{\pi}{4} - \frac{2}{3} \\
&\approx 0.119
\end{aligned}$$



The area of the region is approximately 0.119.
Figure 8.5

Editable Graph

The area represented by the definite integral is shown in Figure 8.5. Try using Simpson's Rule to approximate this integral. With $n = 18$, you should obtain an approximation that is within 0.00001 of the actual value.

Try It | Exploration A | Exploration B

For integrals involving powers of cotangents and cosecants, you can follow a strategy similar to that used for powers of tangents and secants. Also, when integrating trigonometric functions, remember that it sometimes helps to convert the entire integrand to powers of sines and cosines.

EXAMPLE 7 Converting to Sines and Cosines

Find $\displaystyle\int \frac{\sec x}{\tan^2 x}\,dx$.

Solution Because the first four guidelines on page 537 do not apply, try converting the integrand to sines and cosines. In this case, you are able to integrate the resulting powers of sine and cosine as follows.

$$\begin{aligned}\int \frac{\sec x}{\tan^2 x}\,dx &= \int \left(\frac{1}{\cos x}\right)\left(\frac{\cos x}{\sin x}\right)^2 dx\\ &= \int (\sin x)^{-2}(\cos x)\,dx\\ &= -(\sin x)^{-1} + C\\ &= -\csc x + C\end{aligned}$$

Try It Exploration A

Integrals Involving Sine-Cosine Products with Different Angles

FOR FURTHER INFORMATION To learn more about integrals involving sine-cosine products with different angles, see the article "Integrals of Products of Sine and Cosine with Different Arguments" by Sherrie J. Nicol in *The College Mathematics Journal*.

MathArticle

Integrals involving the products of sines and cosines of two *different* angles occur in many applications. In such instances you can use the following product-to-sum identities.

$$\sin mx \sin nx = \frac{1}{2}(\cos[(m-n)x] - \cos[(m+n)x])$$

$$\sin mx \cos nx = \frac{1}{2}(\sin[(m-n)x] + \sin[(m+n)x])$$

$$\cos mx \cos nx = \frac{1}{2}(\cos[(m-n)x] + \cos[(m+n)x])$$

EXAMPLE 8 Using Product-to-Sum Identities

Find $\displaystyle\int \sin 5x \cos 4x\,dx$.

Solution Considering the second product-to-sum identity above, you can write

$$\begin{aligned}\int \sin 5x \cos 4x\,dx &= \frac{1}{2}\int (\sin x + \sin 9x)\,dx\\ &= \frac{1}{2}\left(-\cos x - \frac{\cos 9x}{9}\right) + C\\ &= -\frac{\cos x}{2} - \frac{\cos 9x}{18} + C.\end{aligned}$$

Try It Exploration A

Section 8.4 Trigonometric Substitution

- **Use trigonometric substitution to solve an integral.**
- **Use integrals to model and solve real-life applications.**

EXPLORATION

Integrating a Radical Function
Up to this point in the text, you have not evaluated the following integral.

$$\int_{-1}^{1} \sqrt{1 - x^2}\, dx$$

From geometry, you should be able to find the exact value of this integral—what is it? Using numerical integration, with Simpson's Rule or the Trapezoidal Rule, you can't be sure of the accuracy of the approximation. Why?

Try finding the exact value using the substitution

$$x = \sin\theta \quad \text{and} \quad dx = \cos\theta\, d\theta.$$

Does your answer agree with the value you obtained using geometry?

Trigonometric Substitution

Now that you can evaluate integrals involving powers of trigonometric functions, you can use **trigonometric substitution** to evaluate integrals involving the radicals

$$\sqrt{a^2 - u^2}, \qquad \sqrt{a^2 + u^2}, \qquad \text{and} \qquad \sqrt{u^2 - a^2}.$$

The objective with trigonometric substitution is to eliminate the radical in the integrand. You do this with the Pythagorean identities

$$\cos^2\theta = 1 - \sin^2\theta, \quad \sec^2\theta = 1 + \tan^2\theta, \quad \text{and} \quad \tan^2\theta = \sec^2\theta - 1.$$

For example, if $a > 0$, let $u = a \sin\theta$, where $-\pi/2 \le \theta \le \pi/2$. Then

$$\begin{aligned}\sqrt{a^2 - u^2} &= \sqrt{a^2 - a^2\sin^2\theta}\\ &= \sqrt{a^2(1 - \sin^2\theta)}\\ &= \sqrt{a^2\cos^2\theta}\\ &= a\cos\theta.\end{aligned}$$

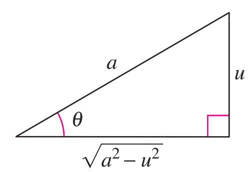
Note that $\cos\theta \ge 0$, because $-\pi/2 \le \theta \le \pi/2$.

Trigonometric Substitution ($a > 0$)

1. For integrals involving $\sqrt{a^2 - u^2}$, let

$$u = a \sin\theta.$$

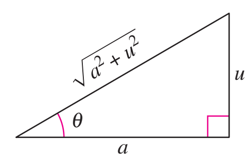
Then $\sqrt{a^2 - u^2} = a\cos\theta$, where $-\pi/2 \le \theta \le \pi/2$.



2. For integrals involving $\sqrt{a^2 + u^2}$, let

$$u = a \tan\theta.$$

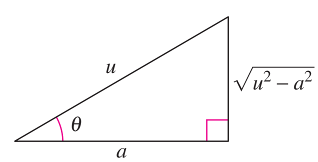
Then $\sqrt{a^2 + u^2} = a\sec\theta$, where $-\pi/2 < \theta < \pi/2$.



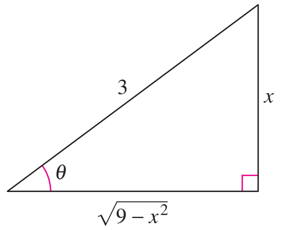
3. For integrals involving $\sqrt{u^2 - a^2}$, let

$$u = a \sec\theta.$$

Then $\sqrt{u^2 - a^2} = \pm a\tan\theta$, where $0 \le \theta < \pi/2$ or $\pi/2 < \theta \le \pi$. Use the positive value if $u > a$ and the negative value if $u < -a$.



NOTE The restrictions on θ ensure that the function that defines the substitution is one-to-one. In fact, these are the same intervals over which the arcsine, arctangent, and arcsecant are defined.



$\sin\theta = \dfrac{x}{3}, \cot\theta = \dfrac{\sqrt{9-x^2}}{x}$

Figure 8.6

EXAMPLE 1 Trigonometric Substitution: $u = a\sin\theta$

Find $\displaystyle\int \frac{dx}{x^2\sqrt{9-x^2}}$.

Solution First, note that none of the basic integration rules applies. To use trigonometric substitution, you should observe that $\sqrt{9-x^2}$ is of the form $\sqrt{a^2-u^2}$. So, you can use the substitution

$$x = a\sin\theta = 3\sin\theta.$$

Using differentiation and the triangle shown in Figure 8.6, you obtain

$$dx = 3\cos\theta\, d\theta, \qquad \sqrt{9-x^2} = 3\cos\theta, \qquad \text{and} \qquad x^2 = 9\sin^2\theta.$$

So, trigonometric substitution yields

$$\begin{aligned}
\int \frac{dx}{x^2\sqrt{9-x^2}} &= \int \frac{3\cos\theta\, d\theta}{(9\sin^2\theta)(3\cos\theta)} && \text{Substitute.}\\
&= \frac{1}{9}\int \frac{d\theta}{\sin^2\theta} && \text{Simplify.}\\
&= \frac{1}{9}\int \csc^2\theta\, d\theta && \text{Trigonometric identity}\\
&= -\frac{1}{9}\cot\theta + C && \text{Apply Cosecant Rule.}\\
&= -\frac{1}{9}\left(\frac{\sqrt{9-x^2}}{x}\right) + C && \text{Substitute for } \cot\theta.\\
&= -\frac{\sqrt{9-x^2}}{9x} + C.
\end{aligned}$$

Note that the triangle in Figure 8.6 can be used to convert the θ's back to x's as follows.

$$\begin{aligned}
\cot\theta &= \frac{\text{adj.}}{\text{opp.}}\\
&= \frac{\sqrt{9-x^2}}{x}
\end{aligned}$$

Try It | Exploration A | Exploration B

TECHNOLOGY Use a computer algebra system to find each definite integral.

$$\int \frac{dx}{\sqrt{9-x^2}} \qquad \int \frac{dx}{x\sqrt{9-x^2}} \qquad \int \frac{dx}{x^2\sqrt{9-x^2}} \qquad \int \frac{dx}{x^3\sqrt{9-x^2}}$$

Then use trigonometric substitution to duplicate the results obtained with the computer algebra system.

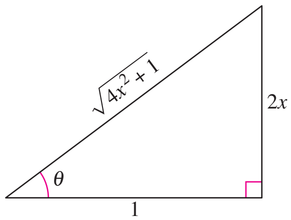
In an earlier chapter, you saw how the inverse hyperbolic functions can be used to evaluate the integrals

$$\int \frac{du}{\sqrt{u^2 \pm a^2}}, \qquad \int \frac{du}{a^2-u^2}, \qquad \text{and} \qquad \int \frac{du}{u\sqrt{a^2 \pm u^2}}.$$

You can also evaluate these integrals using trigonometric substitution. This is shown in the next example.

EXAMPLE 2 Trigonometric Substitution: $u = a \tan \theta$

Find $\displaystyle\int \frac{dx}{\sqrt{4x^2 + 1}}$.



$\tan \theta = 2x$, $\sec \theta = \sqrt{4x^2 + 1}$

Figure 8.7

Solution Let $u = 2x$, $a = 1$, and $2x = \tan \theta$, as shown in Figure 8.7. Then,

$$dx = \frac{1}{2}\sec^2 \theta \, d\theta \qquad \text{and} \qquad \sqrt{4x^2 + 1} = \sec \theta.$$

Trigonometric substitution produces

$$\begin{aligned}
\int \frac{1}{\sqrt{4x^2+1}}\,dx &= \frac{1}{2}\int \frac{\sec^2 \theta \, d\theta}{\sec \theta} && \text{Substitute.}\\
&= \frac{1}{2}\int \sec \theta \, d\theta && \text{Simplify.}\\
&= \frac{1}{2}\ln|\sec \theta + \tan \theta| + C && \text{Apply Secant Rule.}\\
&= \frac{1}{2}\ln\left|\sqrt{4x^2+1} + 2x\right| + C. && \text{Back-substitute.}
\end{aligned}$$

Try checking this result with a computer algebra system. Is the result given in this form or in the form of an inverse hyperbolic function?

Try It

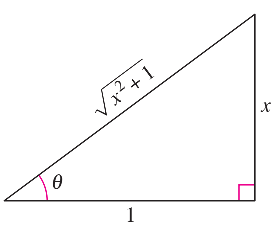
Exploration A Exploration B Exploration C Exploration D

You can extend the use of trigonometric substitution to cover integrals involving expressions such as $(a^2 - u^2)^{n/2}$ by writing the expression as

$$(a^2 - u^2)^{n/2} = \left(\sqrt{a^2 - u^2}\right)^n.$$

EXAMPLE 3 Trigonometric Substitution: Rational Powers

Find $\displaystyle\int \frac{dx}{(x^2 + 1)^{3/2}}$.



$\tan \theta = x$, $\sin \theta = \dfrac{x}{\sqrt{x^2 + 1}}$

Figure 8.8

Solution Begin by writing $(x^2 + 1)^{3/2}$ as $\left(\sqrt{x^2 + 1}\right)^3$. Then, let $a = 1$ and $u = x = \tan \theta$, as shown in Figure 8.8. Using

$$dx = \sec^2 \theta \, d\theta \qquad \text{and} \qquad \sqrt{x^2 + 1} = \sec \theta$$

you can apply trigonometric substitution as follows.

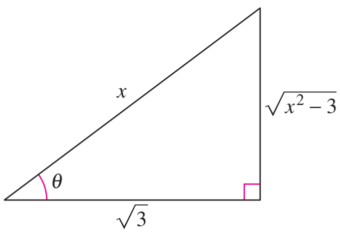
$$\begin{aligned}
\int \frac{dx}{(x^2+1)^{3/2}} &= \int \frac{dx}{\left(\sqrt{x^2+1}\right)^3} && \text{Rewrite denominator.}\\
&= \int \frac{\sec^2 \theta \, d\theta}{\sec^3 \theta} && \text{Substitute.}\\
&= \int \frac{d\theta}{\sec \theta} && \text{Simplify.}\\
&= \int \cos \theta \, d\theta && \text{Trigonometric identity}\\
&= \sin \theta + C && \text{Apply Cosine Rule.}\\
&= \frac{x}{\sqrt{x^2+1}} + C && \text{Back-substitute.}
\end{aligned}$$

Try It Exploration A Exploration B Open Exploration

For definite integrals, it is often convenient to determine the integration limits for θ that avoid converting back to x. You might want to review this procedure in Section 4.5, Examples 8 and 9.

EXAMPLE 4 Converting the Limits of Integration

Evaluate $\displaystyle\int_{\sqrt{3}}^{2} \frac{\sqrt{x^2-3}}{x}\,dx.$



$\sec\theta = \dfrac{x}{\sqrt{3}}$, $\tan\theta = \dfrac{\sqrt{x^2-3}}{\sqrt{3}}$

Figure 8.9

Solution Because $\sqrt{x^2-3}$ has the form $\sqrt{u^2-a^2}$, you can consider

$$u = x, \quad a = \sqrt{3}, \quad \text{and} \quad x = \sqrt{3}\sec\theta$$

as shown in Figure 8.9. Then,

$$dx = \sqrt{3}\sec\theta\tan\theta\,d\theta \qquad \text{and} \qquad \sqrt{x^2-3} = \sqrt{3}\tan\theta.$$

To determine the upper and lower limits of integration, use the substitution $x = \sqrt{3}\sec\theta$, as follows.

Lower Limit	*Upper Limit*
When $x = \sqrt{3}$, $\sec\theta = 1$ and $\theta = 0$.	When $x = 2$, $\sec\theta = \dfrac{2}{\sqrt{3}}$ and $\theta = \dfrac{\pi}{6}$.

So, you have

Integration limits for x | Integration limits for θ

$$\begin{aligned}
\int_{\sqrt{3}}^{2} \frac{\sqrt{x^2-3}}{x}\,dx &= \int_{0}^{\pi/6} \frac{(\sqrt{3}\tan\theta)(\sqrt{3}\sec\theta\tan\theta)\,d\theta}{\sqrt{3}\sec\theta} \\
&= \int_{0}^{\pi/6} \sqrt{3}\tan^2\theta\,d\theta \\
&= \sqrt{3}\int_{0}^{\pi/6} (\sec^2\theta - 1)\,d\theta \\
&= \sqrt{3}\Big[\tan\theta - \theta\Big]_{0}^{\pi/6} \\
&= \sqrt{3}\left(\frac{1}{\sqrt{3}} - \frac{\pi}{6}\right) \\
&= 1 - \frac{\sqrt{3}\pi}{6} \\
&\approx 0.0931.
\end{aligned}$$

Try It | Exploration A | Exploration B

In Example 4, try converting back to the variable x and evaluating the antiderivative at the original limits of integration. You should obtain

$$\int_{\sqrt{3}}^{2} \frac{\sqrt{x^2-3}}{x}\,dx = \sqrt{3}\left[\frac{\sqrt{x^2-3}}{\sqrt{3}} - \operatorname{arcsec}\frac{x}{\sqrt{3}}\right]_{\sqrt{3}}^{2}.$$

When using trigonometric substitution to evaluate definite integrals, you must be careful to check that the values of θ lie in the intervals discussed at the beginning of this section. For instance, if in Example 4 you had been asked to evaluate the definite integral

$$\int_{-2}^{-\sqrt{3}} \frac{\sqrt{x^2-3}}{x}\,dx$$

then using $u = x$ and $a = \sqrt{3}$ in the interval $\left[-2, -\sqrt{3}\right]$ would imply that $u < -a$. So, when determining the upper and lower limits of integration, you would have to choose θ such that $\pi/2 < \theta \le \pi$. In this case the integral would be evaluated as follows.

$$\begin{aligned}
\int_{-2}^{-\sqrt{3}} \frac{\sqrt{x^2-3}}{x}\,dx &= \int_{5\pi/6}^{\pi} \frac{\left(-\sqrt{3}\tan\theta\right)\left(\sqrt{3}\sec\theta\tan\theta\right)d\theta}{\sqrt{3}\sec\theta} \\
&= \int_{5\pi/6}^{\pi} -\sqrt{3}\tan^2\theta\,d\theta \\
&= -\sqrt{3}\int_{5\pi/6}^{\pi} (\sec^2\theta - 1)\,d\theta \\
&= -\sqrt{3}\Big[\tan\theta - \theta\Big]_{5\pi/6}^{\pi} \\
&= -\sqrt{3}\left[(0-\pi) - \left(-\frac{1}{\sqrt{3}} - \frac{5\pi}{6}\right)\right] \\
&= -1 + \frac{\sqrt{3}\pi}{6} \\
&\approx -0.0931
\end{aligned}$$

Trigonometric substitution can be used with completing the square. For instance, try evaluating the following integral.

$$\int \sqrt{x^2-2x}\,dx$$

To begin, you could complete the square and write the integral as

$$\int \sqrt{(x-1)^2-1^2}\,dx.$$

Trigonometric substitution can be used to evaluate the three integrals listed in the following theorem. These integrals will be encountered several times in the remainder of the text. When this happens, we will simply refer to this theorem. (In Exercise 85, you are asked to verify the formulas given in the theorem.)

THEOREM 8.2 Special Integration Formulas ($a > 0$)

1. $\displaystyle\int \sqrt{a^2-u^2}\,du = \frac{1}{2}\left(a^2 \arcsin\frac{u}{a} + u\sqrt{a^2-u^2}\right) + C$

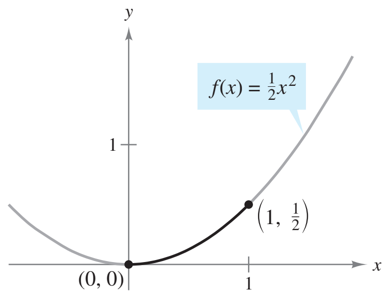
2. $\displaystyle\int \sqrt{u^2-a^2}\,du = \frac{1}{2}\left(u\sqrt{u^2-a^2} - a^2\ln\left|u + \sqrt{u^2-a^2}\right|\right) + C, \quad u > a$

3. $\displaystyle\int \sqrt{u^2+a^2}\,du = \frac{1}{2}\left(u\sqrt{u^2+a^2} + a^2\ln\left|u + \sqrt{u^2+a^2}\right|\right) + C$

Applications

EXAMPLE 5 Finding Arc Length

Find the arc length of the graph of $f(x) = \frac{1}{2}x^2$ from $x = 0$ to $x = 1$ (see Figure 8.10).



The arc length of the curve from $(0, 0)$ to $\left(1, \frac{1}{2}\right)$
Figure 8.10

Editable Graph

Solution Refer to the arc length formula in Section 7.4.

$$s = \int_0^1 \sqrt{1 + [f'(x)]^2}\, dx \qquad \text{Formula for arc length}$$

$$= \int_0^1 \sqrt{1 + x^2}\, dx \qquad f'(x) = x$$

$$= \int_0^{\pi/4} \sec^3 \theta \, d\theta \qquad \text{Let } a = 1 \text{ and } x = \tan \theta.$$

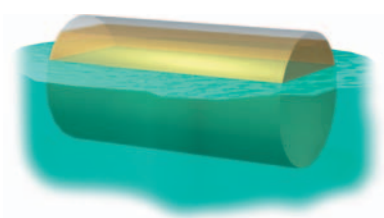
$$= \frac{1}{2}\Big[\sec \theta \tan \theta + \ln|\sec \theta + \tan \theta|\Big]_0^{\pi/4} \qquad \text{Example 5, Section 8.2}$$

$$= \frac{1}{2}\left[\sqrt{2} + \ln\left(\sqrt{2} + 1\right)\right] \approx 1.148$$

Try It Exploration A

EXAMPLE 6 Comparing Two Fluid Forces

A sealed barrel of oil (weighing 48 pounds per cubic foot) is floating in seawater (weighing 64 pounds per cubic foot), as shown in Figures 8.11 and 8.12. (The barrel is not completely full of oil—on its side, the top 0.2 foot of the barrel is empty.) Compare the fluid forces against one end of the barrel from the inside and from the outside.



The barrel is not quite full of oil—the top 0.2 foot of the barrel is empty.
Figure 8.11

Rotatable Graph

Solution In Figure 8.12, locate the coordinate system with the origin at the center of the circle given by $x^2 + y^2 = 1$. To find the fluid force against an end of the barrel *from the inside*, integrate between -1 and 0.8 (using a weight of $w = 48$).

$$F = w\int_c^d h(y)L(y)\, dy \qquad \text{General equation (see Section 7.7)}$$

$$F_{\text{inside}} = 48\int_{-1}^{0.8} (0.8 - y)(2)\sqrt{1 - y^2}\, dy$$

$$= 76.8\int_{-1}^{0.8} \sqrt{1 - y^2}\, dy - 96\int_{-1}^{0.8} y\sqrt{1 - y^2}\, dy$$

To find the fluid force *from the outside*, integrate between -1 and 0.4 (using a weight of $w = 64$).

$$F_{\text{outside}} = 64\int_{-1}^{0.4} (0.4 - y)(2)\sqrt{1 - y^2}\, dy$$

$$= 51.2\int_{-1}^{0.4} \sqrt{1 - y^2}\, dy - 128\int_{-1}^{0.4} y\sqrt{1 - y^2}\, dy$$

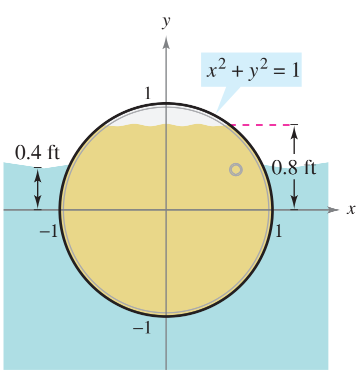


Figure 8.12

The details of integration are left for you to complete in Exercise 84. Intuitively, would you say that the force from the oil (the inside) or the force from the seawater (the outside) is greater? By evaluating these two integrals, you can determine that

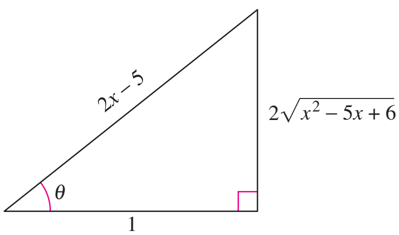
$$F_{\text{inside}} \approx 121.3 \text{ pounds} \qquad \text{and} \qquad F_{\text{outside}} \approx 93.0 \text{ pounds}.$$

Try It Exploration A

Section 8.5 Partial Fractions

- **Understand the concept of a partial fraction decomposition.**
- **Use partial fraction decomposition with linear factors to integrate rational functions.**
- **Use partial fraction decomposition with quadratic factors to integrate rational functions.**

Partial Fractions



$\sec\theta = 2x - 5$
Figure 8.13

This section examines a procedure for decomposing a rational function into simpler rational functions to which you can apply the basic integration formulas. This procedure is called the **method of partial fractions.** To see the benefit of the method of partial fractions, consider the integral

$$\int \frac{1}{x^2 - 5x + 6}\,dx.$$

To evaluate this integral *without* partial fractions, you can complete the square and use trigonometric substitution (see Figure 8.13) to obtain

$$\begin{aligned}
\int \frac{1}{x^2 - 5x + 6}\,dx &= \int \frac{dx}{(x - 5/2)^2 - (1/2)^2} && a = \tfrac{1}{2},\ x - \tfrac{5}{2} = \tfrac{1}{2}\sec\theta \\
&= \int \frac{(1/2)\sec\theta\tan\theta\,d\theta}{(1/4)\tan^2\theta} && dx = \tfrac{1}{2}\sec\theta\tan\theta\,d\theta \\
&= 2\int \csc\theta\,d\theta \\
&= 2\ln|\csc\theta - \cot\theta| + C \\
&= 2\ln\left|\frac{2x-5}{2\sqrt{x^2-5x+6}} - \frac{1}{2\sqrt{x^2-5x+6}}\right| + C \\
&= 2\ln\left|\frac{x-3}{\sqrt{x^2-5x+6}}\right| + C \\
&= 2\ln\left|\frac{\sqrt{x-3}}{\sqrt{x-2}}\right| + C \\
&= \ln\left|\frac{x-3}{x-2}\right| + C \\
&= \ln|x-3| - \ln|x-2| + C.
\end{aligned}$$

Now, suppose you had observed that

$$\frac{1}{x^2 - 5x + 6} = \frac{1}{x-3} - \frac{1}{x-2}. \qquad \text{Partial fraction decomposition}$$

Then you could evaluate the integral easily, as follows.

$$\begin{aligned}
\int \frac{1}{x^2 - 5x + 6}\,dx &= \int \left(\frac{1}{x-3} - \frac{1}{x-2}\right) dx \\
&= \ln|x-3| - \ln|x-2| + C
\end{aligned}$$

This method is clearly preferable to trigonometric substitution. However, its use depends on the ability to factor the denominator, $x^2 - 5x + 6$, and to find the **partial fractions**

$$\frac{1}{x-3} \quad \text{and} \quad -\frac{1}{x-2}.$$

In this section, you will study techniques for finding partial fraction decompositions.

JOHN BERNOULLI (1667–1748)

The method of partial fractions was introduced by John Bernoulli, a Swiss mathematician who was instrumental in the early development of calculus. John Bernoulli was a professor at the University of Basel and taught many outstanding students, the most famous of whom was Leonhard Euler.

MathBio

STUDY TIP In precalculus you learned how to combine functions such as

$$\frac{1}{x-2} + \frac{-1}{x+3} = \frac{5}{(x-2)(x+3)}.$$

The method of partial fractions shows you how to reverse this process.

$$\frac{5}{(x-2)(x+3)} = \frac{?}{x-2} + \frac{?}{x+3}$$

Recall from algebra that every polynomial with real coefficients can be factored into linear and irreducible quadratic factors.* For instance, the polynomial

$$x^5 + x^4 - x - 1$$

can be written as

$$\begin{aligned} x^5 + x^4 - x - 1 &= x^4(x+1) - (x+1) \\ &= (x^4 - 1)(x+1) \\ &= (x^2+1)(x^2-1)(x+1) \\ &= (x^2+1)(x+1)(x-1)(x+1) \\ &= (x-1)(x+1)^2(x^2+1) \end{aligned}$$

where $(x - 1)$ is a linear factor, $(x + 1)^2$ is a repeated linear factor, and $(x^2 + 1)$ is an irreducible quadratic factor. Using this factorization, you can write the partial fraction decomposition of the rational expression

$$\frac{N(x)}{x^5 + x^4 - x - 1}$$

where $N(x)$ is a polynomial of degree less than 5, as follows.

$$\frac{N(x)}{(x-1)(x+1)^2(x^2+1)} = \frac{A}{x-1} + \frac{B}{x+1} + \frac{C}{(x+1)^2} + \frac{Dx+E}{x^2+1}$$

Decomposition of $N(x)/D(x)$ into Partial Fractions

1. **Divide if improper:** If $N(x)/D(x)$ is an improper fraction (that is, if the degree of the numerator is greater than or equal to the degree of the denominator), divide the denominator into the numerator to obtain

$$\frac{N(x)}{D(x)} = (\text{a polynomial}) + \frac{N_1(x)}{D(x)}$$

where the degree of $N_1(x)$ is less than the degree of $D(x)$. Then apply Steps 2, 3, and 4 to the proper rational expression $N_1(x)/D(x)$.

2. **Factor denominator:** Completely factor the denominator into factors of the form

$$(px + q)^m \qquad \text{and} \qquad (ax^2 + bx + c)^n$$

where $ax^2 + bx + c$ is irreducible.

3. **Linear factors:** For each factor of the form $(px + q)^m$, the partial fraction decomposition must include the following sum of m fractions.

$$\frac{A_1}{(px+q)} + \frac{A_2}{(px+q)^2} + \cdots + \frac{A_m}{(px+q)^m}$$

4. **Quadratic factors:** For each factor of the form $(ax^2 + bx + c)^n$, the partial fraction decomposition must include the following sum of n fractions.

$$\frac{B_1x + C_1}{ax^2 + bx + c} + \frac{B_2x + C_2}{(ax^2 + bx + c)^2} + \cdots + \frac{B_nx + C_n}{(ax^2 + bx + c)^n}$$

---

* *For a review of factorization techniques, see* Precalculus, *6th edition, by Larson and Hostetler or* Precalculus: A Graphing Approach, *4th edition, by Larson, Hostetler, and Edwards (Boston, Massachusetts: Houghton Mifflin, 2004 and 2005, respectively).*

Linear Factors

Algebraic techniques for determining the constants in the numerators of a partial decomposition with linear or repeated linear factors are shown in Examples 1 and 2.

EXAMPLE 1 Distinct Linear Factors

Write the partial fraction decomposition for $\dfrac{1}{x^2 - 5x + 6}$.

Solution Because $x^2 - 5x + 6 = (x - 3)(x - 2)$, you should include one partial fraction for each factor and write

$$\frac{1}{x^2 - 5x + 6} = \frac{A}{x - 3} + \frac{B}{x - 2}$$

where A and B are to be determined. Multiplying this equation by the least common denominator $(x - 3)(x - 2)$ yields the **basic equation**

$$1 = A(x - 2) + B(x - 3). \qquad \text{Basic equation}$$

Because this equation is to be true for all x, you can substitute any *convenient* values for x to obtain equations in A and B. The most convenient values are the ones that make particular factors equal to 0.

NOTE Note that the substitutions for x in Example 1 are chosen for their convenience in determining values for A and B; $x = 2$ is chosen to eliminate the term $A(x - 2)$, and $x = 3$ is chosen to eliminate the term $B(x - 3)$. The goal is to make *convenient* substitutions whenever possible.

To solve for A, let $x = 3$ and obtain

$$1 = A(3 - 2) + B(3 - 3) \qquad \text{Let } x = 3 \text{ in basic equation.}$$
$$1 = A(1) + B(0)$$
$$A = 1.$$

To solve for B, let $x = 2$ and obtain

$$1 = A(2 - 2) + B(2 - 3) \qquad \text{Let } x = 2 \text{ in basic equation.}$$
$$1 = A(0) + B(-1)$$
$$B = -1.$$

So, the decomposition is

$$\frac{1}{x^2 - 5x + 6} = \frac{1}{x - 3} - \frac{1}{x - 2}$$

as shown at the beginning of this section.

Try It | Exploration A | Exploration B | Exploration C

Be sure you see that the method of partial fractions is practical only for integrals of rational functions whose denominators factor "nicely." For instance, if the denominator in Example 1 were changed to $x^2 - 5x + 5$, its factorization as

$$x^2 - 5x + 5 = \left[x + \frac{5 + \sqrt{5}}{2}\right]\left[x - \frac{5 - \sqrt{5}}{2}\right]$$

would be too cumbersome to use with partial fractions. In such cases, you should use completing the square or a computer algebra system to perform the integration. If you do this, you should obtain

$$\int \frac{1}{x^2 - 5x + 5}\,dx = \frac{\sqrt{5}}{5}\ln\left|2x - \sqrt{5} - 5\right| - \frac{\sqrt{5}}{5}\ln\left|2x + \sqrt{5} - 5\right| + C.$$

EXAMPLE 2 Repeated Linear Factors

Find $\displaystyle\int \frac{5x^2 + 20x + 6}{x^3 + 2x^2 + x}\,dx.$

Solution Because

$$\begin{aligned} x^3 + 2x^2 + x &= x(x^2 + 2x + 1) \\ &= x(x + 1)^2 \end{aligned}$$

you should include one fraction for *each power* of x and $(x + 1)$ and write

$$\frac{5x^2 + 20x + 6}{x(x + 1)^2} = \frac{A}{x} + \frac{B}{x + 1} + \frac{C}{(x + 1)^2}.$$

Multiplying by the least common denominator $x(x + 1)^2$ yields the *basic equation*

$$5x^2 + 20x + 6 = A(x + 1)^2 + Bx(x + 1) + Cx. \qquad \text{Basic equation}$$

To solve for A, let $x = 0$. This eliminates the B and C terms and yields

$$\begin{aligned} 6 &= A(1) + 0 + 0 \\ A &= 6. \end{aligned}$$

To solve for C, let $x = -1$. This eliminates the A and B terms and yields

$$\begin{aligned} 5 - 20 + 6 &= 0 + 0 - C \\ C &= 9. \end{aligned}$$

The most convenient choices for x have been used, so to find the value of B, you can use *any other value* of x along with the calculated values of A and C. Using $x = 1$, $A = 6$, and $C = 9$ produces

$$\begin{aligned} 5 + 20 + 6 &= A(4) + B(2) + C \\ 31 &= 6(4) + 2B + 9 \\ -2 &= 2B \\ B &= -1. \end{aligned}$$

So, it follows that

$$\begin{aligned} \int \frac{5x^2 + 20x + 6}{x(x + 1)^2}\,dx &= \int \left(\frac{6}{x} - \frac{1}{x + 1} + \frac{9}{(x + 1)^2}\right) dx \\ &= 6 \ln|x| - \ln|x + 1| + 9\frac{(x + 1)^{-1}}{-1} + C \\ &= \ln\left|\frac{x^6}{x + 1}\right| - \frac{9}{x + 1} + C. \end{aligned}$$

Try checking this result by differentiating. Include algebra in your check, simplifying the derivative until you have obtained the original integrand.

Try It | Exploration A | Exploration B | Exploration C

TECHNOLOGY Most computer algebra systems, such as *Derive*, *Maple*, *Mathcad*, *Mathematica*, and the *TI-89*, can be used to convert a rational function to its partial fraction decomposition. For instance, using *Maple*, you obtain the following.

$> \text{convert}\left(\dfrac{5x^2 + 20x + 6}{x^3 + 2x^2 + x}, \text{parfrac}, x\right)$

$\dfrac{6}{x} + \dfrac{9}{(x + 1)^2} - \dfrac{1}{x + 1}$

NOTE It is necessary to make as many substitutions for x as there are unknowns $(A, B, C, \ldots)$ to be determined. For instance, in Example 2, three substitutions $(x = 0,$ $x = -1$, and $x = 1)$ were made to solve for A, B, and C.

Quadratic Factors

When using the method of partial fractions with *linear* factors, a convenient choice of x immediately yields a value for one of the coefficients. With *quadratic* factors, a system of linear equations usually has to be solved, regardless of the choice of x.

EXAMPLE 3 Distinct Linear and Quadratic Factors

Find $\displaystyle\int \frac{2x^3 - 4x - 8}{(x^2 - x)(x^2 + 4)}\,dx.$

Solution Because

$$(x^2 - x)(x^2 + 4) = x(x - 1)(x^2 + 4)$$

you should include one partial fraction for each factor and write

$$\frac{2x^3 - 4x - 8}{x(x - 1)(x^2 + 4)} = \frac{A}{x} + \frac{B}{x - 1} + \frac{Cx + D}{x^2 + 4}.$$

Multiplying by the least common denominator $x(x - 1)(x^2 + 4)$ yields the *basic equation*

$$2x^3 - 4x - 8 = A(x - 1)(x^2 + 4) + Bx(x^2 + 4) + (Cx + D)(x)(x - 1).$$

To solve for A, let $x = 0$ and obtain

$$-8 = A(-1)(4) + 0 + 0 \quad \Rightarrow \quad 2 = A.$$

To solve for B, let $x = 1$ and obtain

$$-10 = 0 + B(5) + 0 \quad \Rightarrow \quad -2 = B.$$

At this point, C and D are yet to be determined. You can find these remaining constants by choosing two other values for x and solving the resulting system of linear equations. If $x = -1$, then, using $A = 2$ and $B = -2$, you can write

$$\begin{aligned} -6 &= (2)(-2)(5) + (-2)(-1)(5) + (-C + D)(-1)(-2) \\ 2 &= -C + D. \end{aligned}$$

If $x = 2$, you have

$$\begin{aligned} 0 &= (2)(1)(8) + (-2)(2)(8) + (2C + D)(2)(1) \\ 8 &= 2C + D. \end{aligned}$$

Solving the linear system by subtracting the first equation from the second

$$\begin{aligned} -C + D &= 2 \\ 2C + D &= 8 \end{aligned}$$

yields $C = 2$. Consequently, $D = 4$, and it follows that

$$\begin{aligned} \int \frac{2x^3 - 4x - 8}{x(x - 1)(x^2 + 4)}\,dx &= \int \left(\frac{2}{x} - \frac{2}{x - 1} + \frac{2x}{x^2 + 4} + \frac{4}{x^2 + 4}\right) dx \\ &= 2\ln|x| - 2\ln|x - 1| + \ln(x^2 + 4) + 2\arctan\frac{x}{2} + C. \end{aligned}$$

Try It | Exploration A | Open Exploration

In Examples 1, 2, and 3, the solution of the basic equation began with substituting values of x that made the linear factors equal to 0. This method works well when the partial fraction decomposition involves linear factors. However, if the decomposition involves only quadratic factors, an alternative procedure is often more convenient.

EXAMPLE 4 Repeated Quadratic Factors

Find $\displaystyle\int \frac{8x^3 + 13x}{(x^2 + 2)^2}\, dx.$

Solution Include one partial fraction for each power of $(x^2 + 2)$ and write

$$\frac{8x^3 + 13x}{(x^2 + 2)^2} = \frac{Ax + B}{x^2 + 2} + \frac{Cx + D}{(x^2 + 2)^2}.$$

Multiplying by the least common denominator $(x^2 + 2)^2$ yields the *basic equation*

$$8x^3 + 13x = (Ax + B)(x^2 + 2) + Cx + D.$$

Expanding the basic equation and collecting like terms produces

$$8x^3 + 13x = Ax^3 + 2Ax + Bx^2 + 2B + Cx + D$$
$$8x^3 + 13x = Ax^3 + Bx^2 + (2A + C)x + (2B + D).$$

Now, you can equate the coefficients of like terms on opposite sides of the equation.

$8 = A$ $\qquad$ $0 = 2B + D$

$$8x^3 + 0x^2 + 13x + 0 = Ax^3 + Bx^2 + (2A + C)x + (2B + D)$$

$0 = B$

$13 = 2A + C$

Using the known values $A = 8$ and $B = 0$, you can write

$$13 = 2A + C = 2(8) + C \quad \Rightarrow \quad C = -3$$
$$0 = 2B + D = 2(0) + D \quad \Rightarrow \quad D = 0.$$

Finally, you can conclude that

$$\int \frac{8x^3 + 13x}{(x^2 + 2)^2}\, dx = \int \left(\frac{8x}{x^2 + 2} + \frac{-3x}{(x^2 + 2)^2}\right) dx$$
$$= 4\ln(x^2 + 2) + \frac{3}{2(x^2 + 2)} + C.$$

Try It | Exploration A | Technology

TECHNOLOGY Use a computer algebra system to evaluate the integral in Example 4—you might find that the form of the antiderivative is different. For instance, when you use a computer algebra system to work Example 4, you obtain

$$\int \frac{8x^3 + 13x}{(x^2 + 2)^2}\, dx = \ln(x^8 + 8x^6 + 24x^4 + 32x^2 + 16) + \frac{3}{2(x^2 + 2)} + C.$$

Is this result equivalent to that obtained in Example 4?

When integrating rational expressions, keep in mind that for *improper* rational expressions such as

$$\frac{N(x)}{D(x)} = \frac{2x^3 + x^2 - 7x + 7}{x^2 + x - 2}$$

you must first divide to obtain

$$\frac{N(x)}{D(x)} = 2x - 1 + \frac{-2x + 5}{x^2 + x - 2}.$$

The proper rational expression is then decomposed into its partial fractions by the usual methods. Here are some guidelines for solving the basic equation that is obtained in a partial fraction decomposition.

Guidelines for Solving the Basic Equation

Linear Factors

1. Substitute the roots of the distinct linear factors into the basic equation.
2. For repeated linear factors, use the coefficients determined in guideline 1 to rewrite the basic equation. Then substitute other convenient values of x and solve for the remaining coefficients.

Quadratic Factors

1. Expand the basic equation.
2. Collect terms according to powers of x.
3. Equate the coefficients of like powers to obtain a system of linear equations involving A, B, C, and so on.
4. Solve the system of linear equations.

Before concluding this section, here are a few things you should remember. First, it is not necessary to use the partial fractions technique on all rational functions. For instance, the following integral is evaluated more easily by the Log Rule.

$$\begin{aligned}\int \frac{x^2 + 1}{x^3 + 3x - 4}\,dx &= \frac{1}{3}\int \frac{3x^2 + 3}{x^3 + 3x - 4}\,dx \\ &= \frac{1}{3}\ln|x^3 + 3x - 4| + C\end{aligned}$$

Second, if the integrand is not in reduced form, reducing it may eliminate the need for partial fractions, as shown in the following integral.

$$\begin{aligned}\int \frac{x^2 - x - 2}{x^3 - 2x - 4}\,dx &= \int \frac{(x + 1)(x - 2)}{(x - 2)(x^2 + 2x + 2)}\,dx \\ &= \int \frac{x + 1}{x^2 + 2x + 2}\,dx \\ &= \frac{1}{2}\ln|x^2 + 2x + 2| + C\end{aligned}$$

Finally, partial fractions can be used with some quotients involving transcendental functions. For instance, the substitution $u = \sin x$ allows you to write

$$\int \frac{\cos x}{\sin x(\sin x - 1)}\,dx = \int \frac{du}{u(u - 1)}. \qquad u = \sin x,\ du = \cos x\,dx$$

Section 8.6 Integration by Tables and Other Integration Techniques

- **Evaluate an indefinite integral using a table of integrals.**
- **Evaluate an indefinite integral using reduction formulas.**
- **Evaluate an indefinite integral involving rational functions of sine and cosine.**

Integration by Tables

So far in this chapter you have studied several integration techniques that can be used with the basic integration rules. But merely knowing *how* to use the various techniques is not enough. You also need to know *when* to use them. Integration is first and foremost a problem of recognition. That is, you must recognize which rule or technique to apply to obtain an antiderivative. Frequently, a slight alteration of an integrand will require a different integration technique (or produce a function whose antiderivative is not an elementary function), as shown below.

$$\int x \ln x \, dx = \frac{x^2}{2} \ln x - \frac{x^2}{4} + C \qquad \text{Integration by parts}$$

$$\int \frac{\ln x}{x} \, dx = \frac{(\ln x)^2}{2} + C \qquad \text{Power Rule}$$

$$\int \frac{1}{x \ln x} \, dx = \ln|\ln x| + C \qquad \text{Log Rule}$$

$$\int \frac{x}{\ln x} \, dx = \, ? \qquad \text{Not an elementary function}$$

TECHNOLOGY A computer algebra system consists, in part, of a database of integration formulas. The primary difference between using a computer algebra system and using tables of integrals is that with a computer algebra system the computer searches through the database to find a fit. With integration tables, *you* must do the searching.

Many people find tables of integrals to be a valuable supplement to the integration techniques discussed in this chapter. Tables of common integrals can be found in Appendix B. **Integration by tables** is not a "cure-all" for all of the difficulties that can accompany integration—using tables of integrals requires considerable thought and insight and often involves substitution.

Each integration formula in Appendix B can be developed using one or more of the techniques in this chapter. You should try to verify several of the formulas. For instance, Formula 4

$$\int \frac{u}{(a + bu)^2} \, du = \frac{1}{b^2} \left(\frac{a}{a + bu} + \ln|a + bu| \right) + C \qquad \text{Formula 4}$$

can be verified using the method of partial fractions, and Formula 19

$$\int \frac{\sqrt{a + bu}}{u} \, du = 2\sqrt{a + bu} + a \int \frac{du}{u\sqrt{a + bu}} \qquad \text{Formula 19}$$

can be verified using integration by parts. Note that the integrals in Appendix B are classified according to forms involving the following.

u^n	$(a + bu)$
$(a + bu + cu^2)$	$\sqrt{a + bu}$
$(a^2 \pm u^2)$	$\sqrt{u^2 \pm a^2}$
$\sqrt{a^2 - u^2}$	Trigonometric functions
Inverse trigonometric functions	Exponential functions
Logarithmic functions	

EXPLORATION

Use the tables of integrals in Appendix B and the substitution

$$u = \sqrt{x-1}$$

to evaluate the integral in Example 1. If you do this, you should obtain

$$\int \frac{dx}{x\sqrt{x-1}} = \int \frac{2\,du}{u^2+1}.$$

Does this produce the same result as that obtained in Example 1?

EXAMPLE 1 Integration by Tables

Find $\displaystyle\int \frac{dx}{x\sqrt{x-1}}$.

Solution Because the expression inside the radical is linear, you should consider forms involving $\sqrt{a+bu}$.

$$\int \frac{du}{u\sqrt{a+bu}} = \frac{2}{\sqrt{-a}} \arctan \sqrt{\frac{a+bu}{-a}} + C \qquad \text{Formula 17 } (a < 0)$$

Let $a = -1$, $b = 1$, and $u = x$. Then $du = dx$, and you can write

$$\int \frac{dx}{x\sqrt{x-1}} = 2 \arctan \sqrt{x-1} + C.$$

Try It Exploration A

EXAMPLE 2 Integration by Tables

Find $\displaystyle\int x\sqrt{x^4-9}\,dx$.

Solution Because the radical has the form $\sqrt{u^2-a^2}$, you should consider Formula 26.

$$\int \sqrt{u^2-a^2}\,du = \frac{1}{2}\left(u\sqrt{u^2-a^2} - a^2 \ln\left|u + \sqrt{u^2-a^2}\right|\right) + C$$

Let $u = x^2$ and $a = 3$. Then $du = 2x\,dx$, and you have

$$\begin{aligned}\int x\sqrt{x^4-9}\,dx &= \frac{1}{2}\int \sqrt{(x^2)^2-3^2}\,(2x)\,dx\\ &= \frac{1}{4}\left(x^2\sqrt{x^4-9} - 9\ln\left|x^2+\sqrt{x^4-9}\right|\right) + C.\end{aligned}$$

Try It Exploration A Open Exploration

EXAMPLE 3 Integration by Tables

Find $\displaystyle\int \frac{x}{1+e^{-x^2}}\,dx$.

Solution Of the forms involving e^u, consider the following formula.

$$\int \frac{du}{1+e^u} = u - \ln(1+e^u) + C \qquad \text{Formula 84}$$

Let $u = -x^2$. Then $du = -2x\,dx$, and you have

$$\begin{aligned}\int \frac{x}{1+e^{-x^2}}\,dx &= -\frac{1}{2}\int \frac{-2x\,dx}{1+e^{-x^2}}\\ &= -\frac{1}{2}\left[-x^2 - \ln(1+e^{-x^2})\right] + C\\ &= \frac{1}{2}\left[x^2 + \ln(1+e^{-x^2})\right] + C.\end{aligned}$$

Try It Exploration A

TECHNOLOGY Example 3 shows the importance of having several solution techniques at your disposal. This integral is not difficult to solve with a table, but when it was entered into a well-known computer algebra system, the utility was unable to find the antiderivative.

Reduction Formulas

Several of the integrals in the integration tables have the form $\int f(x)\,dx = g(x) + \int h(x)\,dx$. Such integration formulas are called **reduction formulas** because they reduce a given integral to the sum of a function and a simpler integral.

EXAMPLE 4 Using a Reduction Formula

Find $\displaystyle\int x^3 \sin x\,dx$.

Solution Consider the following three formulas.

$$\int u \sin u\,du = \sin u - u\cos u + C \qquad \text{Formula 52}$$

$$\int u^n \sin u\,du = -u^n \cos u + n\int u^{n-1}\cos u\,du \qquad \text{Formula 54}$$

$$\int u^n \cos u\,du = u^n \sin u - n\int u^{n-1}\sin u\,du \qquad \text{Formula 55}$$

Using Formula 54, Formula 55, and then Formula 52 produces

$$\begin{aligned}\int x^3 \sin x\,dx &= -x^3\cos x + 3\int x^2\cos x\,dx\\ &= -x^3\cos x + 3\left(x^2\sin x - 2\int x\sin x\,dx\right)\\ &= -x^3\cos x + 3x^2\sin x + 6x\cos x - 6\sin x + C.\end{aligned}$$

Try It Exploration A

TECHNOLOGY Sometimes when you use computer algebra systems you obtain results that look very different, but are actually equivalent. Here is how several different systems evaluated the integral in Example 5.

Maple

$\sqrt{3-5x} - \sqrt{3}\,\operatorname{arctanh}\left(\frac{1}{3}\sqrt{3-5x}\sqrt{3}\right)$

Derive

$\sqrt{3}\ln\left[\dfrac{\sqrt{(3-5x)}-\sqrt{3}}{\sqrt{x}}\right] + \sqrt{(3-5x)}$

Mathematica

Sqrt[3 − 5x] −

Sqrt[3] ArcTanh$\left[\dfrac{\text{Sqrt}[3-5x]}{\text{Sqrt}[3]}\right]$

Mathcad

$\sqrt{3-5x} + \frac{1}{2}\sqrt{3}\ln\left[-\frac{1}{5}\dfrac{\left(-6+5x+2\sqrt{3}\sqrt{3-5x}\right)}{x}\right]$

Notice that computer algebra systems do not include a constant of integration.

EXAMPLE 5 Using a Reduction Formula

Find $\displaystyle\int \frac{\sqrt{3-5x}}{2x}\,dx$.

Solution Consider the following two formulas.

$$\int \frac{du}{u\sqrt{a+bu}} = \frac{1}{\sqrt{a}}\ln\left|\frac{\sqrt{a+bu}-\sqrt{a}}{\sqrt{a+bu}+\sqrt{a}}\right| + C \qquad \text{Formula 17 } (a > 0)$$

$$\int \frac{\sqrt{a+bu}}{u}\,du = 2\sqrt{a+bu} + a\int \frac{du}{u\sqrt{a+bu}} \qquad \text{Formula 19}$$

Using Formula 19, with $a = 3$, $b = -5$, and $u = x$, produces

$$\begin{aligned}\frac{1}{2}\int \frac{\sqrt{3-5x}}{x}\,dx &= \frac{1}{2}\left(2\sqrt{3-5x} + 3\int \frac{dx}{x\sqrt{3-5x}}\right)\\ &= \sqrt{3-5x} + \frac{3}{2}\int \frac{dx}{x\sqrt{3-5x}}.\end{aligned}$$

Using Formula 17, with $a = 3$, $b = -5$, and $u = x$, produces

$$\begin{aligned}\int \frac{\sqrt{3-5x}}{2x}\,dx &= \sqrt{3-5x} + \frac{3}{2}\left(\frac{1}{\sqrt{3}}\ln\left|\frac{\sqrt{3-5x}-\sqrt{3}}{\sqrt{3-5x}+\sqrt{3}}\right|\right) + C\\ &= \sqrt{3-5x} + \frac{\sqrt{3}}{2}\ln\left|\frac{\sqrt{3-5x}-\sqrt{3}}{\sqrt{3-5x}+\sqrt{3}}\right| + C.\end{aligned}$$

Try It Exploration A Exploration B

Rational Functions of Sine and Cosine

EXAMPLE 6 Integration by Tables

Find $\displaystyle\int \frac{\sin 2x}{2 + \cos x}\,dx$.

Solution Substituting $2 \sin x \cos x$ for $\sin 2x$ produces

$$\int \frac{\sin 2x}{2 + \cos x}\,dx = 2\int \frac{\sin x \cos x}{2 + \cos x}\,dx.$$

A check of the forms involving $\sin u$ or $\cos u$ in Appendix B shows that none of those listed applies. So, you can consider forms involving $a + bu$. For example,

$$\int \frac{u\,du}{a + bu} = \frac{1}{b^2}(bu - a \ln|a + bu|) + C. \qquad \text{Formula 3}$$

Let $a = 2$, $b = 1$, and $u = \cos x$. Then $du = -\sin x\,dx$, and you have

$$\begin{aligned} 2\int \frac{\sin x \cos x}{2 + \cos x}\,dx &= -2\int \frac{\cos x(-\sin x\,dx)}{2 + \cos x} \\ &= -2(\cos x - 2\ln|2 + \cos x|) + C \\ &= -2\cos x + 4\ln|2 + \cos x| + C. \end{aligned}$$

Try It Exploration A

Example 6 involves a rational expression of $\sin x$ and $\cos x$. If you are unable to find an integral of this form in the integration tables, try using the following special substitution to convert the trigonometric expression to a standard rational expression.

Substitution for Rational Functions of Sine and Cosine

For integrals involving rational functions of sine and cosine, the substitution

$$u = \frac{\sin x}{1 + \cos x} = \tan\frac{x}{2}$$

yields

$$\cos x = \frac{1 - u^2}{1 + u^2}, \quad \sin x = \frac{2u}{1 + u^2}, \quad \text{and} \quad dx = \frac{2\,du}{1 + u^2}.$$

Proof From the substitution for u, it follows that

$$u^2 = \frac{\sin^2 x}{(1 + \cos x)^2} = \frac{1 - \cos^2 x}{(1 + \cos x)^2} = \frac{1 - \cos x}{1 + \cos x}.$$

Solving for $\cos x$ produces $\cos x = (1 - u^2)/(1 + u^2)$. To find $\sin x$, write $u = \sin x/(1 + \cos x)$ as

$$\sin x = u(1 + \cos x) = u\left(1 + \frac{1 - u^2}{1 + u^2}\right) = \frac{2u}{1 + u^2}.$$

Finally, to find dx, consider $u = \tan(x/2)$. Then you have $\arctan u = x/2$ and $dx = (2\,du)/(1 + u^2)$.

Section 8.7 Indeterminate Forms and L'Hôpital's Rule

- **Recognize limits that produce indeterminate forms.**
- **Apply L'Hôpital's Rule to evaluate a limit.**

Indeterminate Forms

Recall from Chapters 1 and 3 that the forms $0/0$ and ∞/∞ are called *indeterminate* because they do not guarantee that a limit exists, nor do they indicate what the limit is, if one does exist. When you encountered one of these indeterminate forms earlier in the text, you attempted to rewrite the expression by using various algebraic techniques.

Indeterminate Form	*Limit*	*Algebraic Technique*
$\dfrac{0}{0}$	$\displaystyle\lim_{x\to-1}\frac{2x^2-2}{x+1}=\lim_{x\to-1}2(x-1)=-4$	Divide numerator and denominator by $(x+1)$.
$\dfrac{\infty}{\infty}$	$\displaystyle\lim_{x\to\infty}\frac{3x^2-1}{2x^2+1}=\lim_{x\to\infty}\frac{3-(1/x^2)}{2+(1/x^2)}=\frac{3}{2}$	Divide numerator and denominator by x^2.

Occasionally, you can extend these algebraic techniques to find limits of transcendental functions. For instance, the limit

$$\lim_{x\to0}\frac{e^{2x}-1}{e^x-1}$$

produces the indeterminate form $0/0$. Factoring and then dividing produces

$$\lim_{x\to0}\frac{e^{2x}-1}{e^x-1}=\lim_{x\to0}\frac{(e^x+1)(e^x-1)}{e^x-1}=\lim_{x\to0}(e^x+1)=2.$$

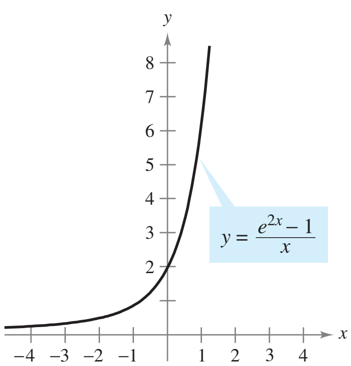
However, not all indeterminate forms can be evaluated by algebraic manipulation. This is often true when *both* algebraic and transcendental functions are involved. For instance, the limit

$$\lim_{x\to0}\frac{e^{2x}-1}{x}$$

produces the indeterminate form $0/0$. Rewriting the expression to obtain

$$\lim_{x\to0}\left(\frac{e^{2x}}{x}-\frac{1}{x}\right)$$

merely produces another indeterminate form, $\infty-\infty$. Of course, you could use technology to estimate the limit, as shown in the table and in Figure 8.14. From the table and the graph, the limit appears to be 2. (This limit will be verified in Example 1.)



The limit as x approaches 0 appears to be 2.
Figure 8.14

x	-1	-0.1	-0.01	-0.001	0	0.001	0.01	0.1	1
$\dfrac{e^{2x}-1}{x}$	0.865	1.813	1.980	1.998	?	2.002	2.020	2.214	6.389

GUILLAUME L'HÔPITAL (1661–1704)

L'Hôpital's Rule is named after the French mathematician Guillaume François Antoine de L'Hôpital. L'Hôpital is credited with writing the first text on differential calculus (in 1696) in which the rule publicly appeared. It was recently discovered that the rule and its proof were written in a letter from John Bernoulli to L'Hôpital. "… I acknowledge that I owe very much to the bright minds of the Bernoulli brothers. … I have made free use of their discoveries …," said L'Hôpital.

MathBio

L'Hôpital's Rule

To find the limit illustrated in Figure 8.14, you can use a theorem called **L'Hôpital's Rule.** This theorem states that under certain conditions the limit of the quotient $f(x)/g(x)$ is determined by the limit of the quotient of the derivatives

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

To prove this theorem, you can use a more general result called the **Extended Mean Value Theorem.**

THEOREM 8.3 The Extended Mean Value Theorem

If f and g are differentiable on an open interval (a, b) and continuous on $[a, b]$ such that $g'(x) \neq 0$ for any x in (a, b), then there exists a point c in (a, b) such that

$$\frac{f'(c)}{g'(c)} = \frac{f(b) - f(a)}{g(b) - g(a)}.$$

NOTE To see why this is called the Extended Mean Value Theorem, consider the special case in which $g(x) = x$. For this case, you obtain the "standard" Mean Value Theorem as presented in Section 3.2.

The Extended Mean Value Theorem and L'Hôpital's Rule are both proved in Appendix A.

THEOREM 8.4 L'Hôpital's Rule

Let f and g be functions that are differentiable on an open interval (a, b) containing c, except possibly at c itself. Assume that $g'(x) \neq 0$ for all x in (a, b), except possibly at c itself. If the limit of $f(x)/g(x)$ as x approaches c produces the indeterminate form $0/0$, then

$$\lim_{x\to c} \frac{f(x)}{g(x)} = \lim_{x\to c} \frac{f'(x)}{g'(x)}$$

provided the limit on the right exists (or is infinite). This result also applies if the limit of $f(x)/g(x)$ as x approaches c produces any one of the indeterminate forms ∞/∞, $(-\infty)/\infty$, $\infty/(-\infty)$, or $(-\infty)/(-\infty)$.

NOTE People occasionally use L'Hôpital's Rule incorrectly by applying the Quotient Rule to $f(x)/g(x)$. Be sure you see that the rule involves $f'(x)/g'(x)$, not the derivative of $f(x)/g(x)$.

FOR FURTHER INFORMATION To enhance your understanding of the necessity of the restriction that $g'(x)$ be nonzero for all x in (a, b), except possibly at c, see the article "Counterexamples to L'Hôpital's Rule" by R. P. Boas in *The American Mathematical Monthly*.

MathArticle

L'Hôpital's Rule can also be applied to one-sided limits. For instance, if the limit of $f(x)/g(x)$ as x approaches c *from the right* produces the indeterminate form $0/0$, then

$$\lim_{x\to c^+} \frac{f(x)}{g(x)} = \lim_{x\to c^+} \frac{f'(x)}{g'(x)}$$

provided the limit exists (or is infinite).

TECHNOLOGY *Numerical and Graphical Approaches* Use a numerical or a graphical approach to approximate each limit.

a. $\lim\limits_{x \to 0} \dfrac{2^{2x} - 1}{x}$

b. $\lim\limits_{x \to 0} \dfrac{3^{2x} - 1}{x}$

c. $\lim\limits_{x \to 0} \dfrac{4^{2x} - 1}{x}$

d. $\lim\limits_{x \to 0} \dfrac{5^{2x} - 1}{x}$

What pattern do you observe? Does an analytic approach have an advantage for these limits? If so, explain your reasoning.

EXAMPLE 1 Indeterminate Form 0/0

Evaluate $\lim\limits_{x \to 0} \dfrac{e^{2x} - 1}{x}$.

Solution Because direct substitution results in the indeterminate form $0/0$

$$\lim_{x \to 0} \frac{e^{2x} - 1}{x} \quad \begin{matrix} \nearrow \lim\limits_{x \to 0} (e^{2x} - 1) = 0 \\ \searrow \lim\limits_{x \to 0} x = 0 \end{matrix}$$

you can apply L'Hôpital's Rule as shown below.

$$\lim_{x \to 0} \frac{e^{2x} - 1}{x} = \lim_{x \to 0} \frac{\dfrac{d}{dx}[e^{2x} - 1]}{\dfrac{d}{dx}[x]} \qquad \text{Apply L'Hôpital's Rule.}$$

$$= \lim_{x \to 0} \frac{2e^{2x}}{1} \qquad \text{Differentiate numerator and denominator.}$$

$$= 2 \qquad \text{Evaluate the limit.}$$

Try It **Exploration A** **Exploration B**

NOTE In writing the string of equations in Example 1, you actually do not know that the first limit is equal to the second until you have shown that the second limit exists. In other words, if the second limit had not existed, it would not have been permissible to apply L'Hôpital's Rule.

Another form of L'Hôpital's Rule states that if the limit of $f(x)/g(x)$ as x approaches ∞ (or $-\infty$) produces the indeterminate form $0/0$ or ∞/∞, then

$$\lim_{x \to \infty} \frac{f(x)}{g(x)} = \lim_{x \to \infty} \frac{f'(x)}{g'(x)}$$

provided the limit on the right exists.

EXAMPLE 2 Indeterminate Form ∞/∞

Evaluate $\lim\limits_{x \to \infty} \dfrac{\ln x}{x}$.

Solution Because direct substitution results in the indeterminate form ∞/∞, you can apply L'Hôpital's Rule to obtain

$$\lim_{x \to \infty} \frac{\ln x}{x} = \lim_{x \to \infty} \frac{\dfrac{d}{dx}[\ln x]}{\dfrac{d}{dx}[x]} \qquad \text{Apply L'Hôpital's Rule.}$$

$$= \lim_{x \to \infty} \frac{1}{x} \qquad \text{Differentiate numerator and denominator.}$$

$$= 0. \qquad \text{Evaluate the limit.}$$

NOTE Try graphing $y_1 = \ln x$ and $y_2 = x$ in the same viewing window. Which function grows faster as x approaches ∞? How is this observation related to Example 2?

Try It **Exploration A** **Exploration B**

Occasionally it is necessary to apply L'Hôpital's Rule more than once to remove an indeterminate form, as shown in Example 3.

EXAMPLE 3 Applying L'Hôpital's Rule More Than Once

Evaluate $\lim\limits_{x \to -\infty} \dfrac{x^2}{e^{-x}}$.

Solution Because direct substitution results in the indeterminate form ∞/∞, you can apply L'Hôpital's Rule.

$$\lim_{x \to -\infty} \frac{x^2}{e^{-x}} = \lim_{x \to -\infty} \frac{\frac{d}{dx}[x^2]}{\frac{d}{dx}[e^{-x}]} = \lim_{x \to -\infty} \frac{2x}{-e^{-x}}$$

This limit yields the indeterminate form $(-\infty)/(-\infty)$, so you can apply L'Hôpital's Rule again to obtain

$$\lim_{x \to -\infty} \frac{2x}{-e^{-x}} = \lim_{x \to -\infty} \frac{\frac{d}{dx}[2x]}{\frac{d}{dx}[-e^{-x}]} = \lim_{x \to -\infty} \frac{2}{e^{-x}} = 0.$$

Try It   Exploration A

In addition to the forms $0/0$ and ∞/∞, there are other indeterminate forms such as $0 \cdot \infty$, 1^∞, ∞^0, 0^0, and $\infty - \infty$. For example, consider the following four limits that lead to the indeterminate form $0 \cdot \infty$.

$$\underbrace{\lim_{x \to 0} (x)\left(\frac{1}{x}\right)}_{\text{Limit is 1.}}, \quad \underbrace{\lim_{x \to 0} (x)\left(\frac{2}{x}\right)}_{\text{Limit is 2.}}, \quad \underbrace{\lim_{x \to \infty} (x)\left(\frac{1}{e^x}\right)}_{\text{Limit is 0.}}, \quad \underbrace{\lim_{x \to \infty} (e^x)\left(\frac{1}{x}\right)}_{\text{Limit is } \infty.}$$

Because each limit is different, it is clear that the form $0 \cdot \infty$ is indeterminate in the sense that it does not determine the value (or even the existence) of the limit. The following examples indicate methods for evaluating these forms. Basically, you attempt to convert each of these forms to $0/0$ or ∞/∞ so that L'Hôpital's Rule can be applied.

EXAMPLE 4 Indeterminate Form $0 \cdot \infty$

Evaluate $\lim\limits_{x \to \infty} e^{-x}\sqrt{x}$.

Solution Because direct substitution produces the indeterminate form $0 \cdot \infty$, you should try to rewrite the limit to fit the form $0/0$ or ∞/∞. In this case, you can rewrite the limit to fit the second form.

$$\lim_{x \to \infty} e^{-x}\sqrt{x} = \lim_{x \to \infty} \frac{\sqrt{x}}{e^x}$$

Now, by L'Hôpital's Rule, you have

$$\lim_{x \to \infty} \frac{\sqrt{x}}{e^x} = \lim_{x \to \infty} \frac{1/(2\sqrt{x})}{e^x} = \lim_{x \to \infty} \frac{1}{2\sqrt{x}\,e^x} = 0.$$

Try It   Exploration A   Exploration B

If rewriting a limit in one of the forms $0/0$ or ∞/∞ does not seem to work, try the other form. For instance, in Example 4 you can write the limit as

$$\lim_{x\to\infty} e^{-x}\sqrt{x} = \lim_{x\to\infty} \frac{e^{-x}}{x^{-1/2}}$$

which yields the indeterminate form $0/0$. As it happens, applying L'Hôpital's Rule to this limit produces

$$\lim_{x\to\infty} \frac{e^{-x}}{x^{-1/2}} = \lim_{x\to\infty} \frac{-e^{-x}}{-1/(2x^{3/2})}$$

which also yields the indeterminate form $0/0$.

The indeterminate forms 1^∞, ∞^0, and 0^0 arise from limits of functions that have variable bases and variable exponents. When you previously encountered this type of function, you used logarithmic differentiation to find the derivative. You can use a similar procedure when taking limits, as shown in the next example.

EXAMPLE 5 Indeterminate Form 1^∞

Evaluate $\displaystyle\lim_{x\to\infty}\left(1+\frac{1}{x}\right)^x$.

Solution Because direct substitution yields the indeterminate form 1^∞, you can proceed as follows. To begin, assume that the limit exists and is equal to y.

$$y = \lim_{x\to\infty}\left(1+\frac{1}{x}\right)^x$$

Taking the natural logarithm of each side produces

$$\ln y = \ln\left[\lim_{x\to\infty}\left(1+\frac{1}{x}\right)^x\right].$$

Because the natural logarithmic function is continuous, you can write

$$\ln y = \lim_{x\to\infty}\left[x\ln\left(1+\frac{1}{x}\right)\right] \qquad \text{Indeterminate form } \infty\cdot 0$$

$$= \lim_{x\to\infty}\left(\frac{\ln[1+(1/x)]}{1/x}\right) \qquad \text{Indeterminate form } 0/0$$

$$= \lim_{x\to\infty}\left(\frac{(-1/x^2)\{1/[1+(1/x)]\}}{-1/x^2}\right) \qquad \text{L'Hôpital's Rule}$$

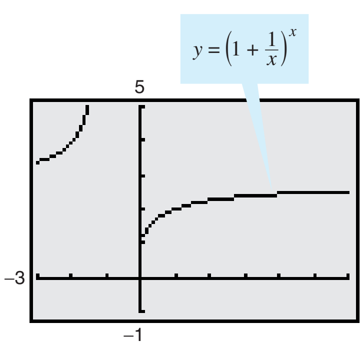
$$= \lim_{x\to\infty}\frac{1}{1+(1/x)}$$

$$= 1.$$

Now, because you have shown that $\ln y = 1$, you can conclude that $y = e$ and obtain

$$\lim_{x\to\infty}\left(1+\frac{1}{x}\right)^x = e.$$

You can use a graphing utility to confirm this result, as shown in Figure 8.15.



The limit of $[1+(1/x)]^x$ as x approaches infinity is e.
Figure 8.15

Editable Graph

Try It

Exploration A

L'Hôpital's Rule can also be applied to one-sided limits, as demonstrated in Examples 6 and 7.

EXAMPLE 6 Indeterminate Form 0^0

Find $\lim_{x \to 0^+} (\sin x)^x$.

Solution Because direct substitution produces the indeterminate form 0^0, you can proceed as shown below. To begin, assume that the limit exists and is equal to y.

$$y = \lim_{x \to 0^+} (\sin x)^x \qquad \text{Indeterminate form } 0^0$$

$$\ln y = \ln\left[\lim_{x \to 0^+} (\sin x)^x\right] \qquad \text{Take natural log of each side.}$$

$$= \lim_{x \to 0^+} \left[\ln(\sin x)^x\right] \qquad \text{Continuity}$$

$$= \lim_{x \to 0^+} \left[x \ln(\sin x)\right] \qquad \text{Indeterminate form } 0 \cdot (-\infty)$$

$$= \lim_{x \to 0^+} \frac{\ln(\sin x)}{1/x} \qquad \text{Indeterminate form } -\infty/\infty$$

$$= \lim_{x \to 0^+} \frac{\cot x}{-1/x^2} \qquad \text{L'Hôpital's Rule}$$

$$= \lim_{x \to 0^+} \frac{-x^2}{\tan x} \qquad \text{Indeterminate form } 0/0$$

$$= \lim_{x \to 0^+} \frac{-2x}{\sec^2 x} = 0 \qquad \text{L'Hôpital's Rule}$$

Now, because $\ln y = 0$, you can conclude that $y = e^0 = 1$, and it follows that

$$\lim_{x \to 0^+} (\sin x)^x = 1.$$

Try It | Exploration A | Open Exploration

TECHNOLOGY When evaluating complicated limits such as the one in Example 6, it is helpful to check the reasonableness of the solution with a computer or with a graphing utility. For instance, the calculations in the following table and the graph in Figure 8.16 are consistent with the conclusion that $(\sin x)^x$ approaches 1 as x approaches 0 from the right.

x	1.0	0.1	0.01	0.001	0.0001	0.00001
$(\sin x)^x$	0.8415	0.7942	0.9550	0.9931	0.9991	0.9999

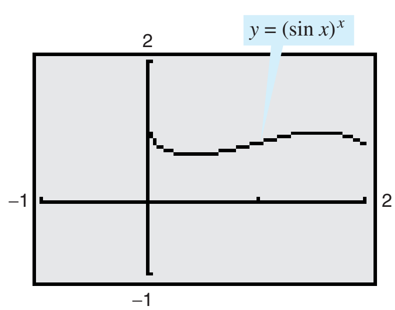
Use a computer algebra system or graphing utility to estimate the following limits:

$$\lim_{x \to 0} (1 - \cos x)^x$$

and

$$\lim_{x \to 0^+} (\tan x)^x.$$

Then see if you can verify your estimates analytically.



The limit of $(\sin x)^x$ is 1 as x approaches 0 from the right.
Figure 8.16

EXAMPLE 7 Indeterminate Form $\infty - \infty$

STUDY TIP In each of the examples presented in this section, L'Hôpital's Rule is used to find a limit that exists. It can also be used to conclude that a limit is infinite. For instance, try using L'Hôpital's Rule to show that

$$\lim_{x\to\infty} \frac{e^x}{x} = \infty.$$

Evaluate $\displaystyle \lim_{x\to 1^+} \left(\frac{1}{\ln x} - \frac{1}{x-1}\right)$.

Solution Because direct substitution yields the indeterminate form $\infty - \infty$, you should try to rewrite the expression to produce a form to which you can apply L'Hôpital's Rule. In this case, you can combine the two fractions to obtain

$$\lim_{x\to 1^+} \left(\frac{1}{\ln x} - \frac{1}{x-1}\right) = \lim_{x\to 1^+} \left[\frac{x-1-\ln x}{(x-1)\ln x}\right].$$

Now, because direct substitution produces the indeterminate form $0/0$, you can apply L'Hôpital's Rule to obtain

$$\begin{aligned}
\lim_{x\to 1^+} \left(\frac{1}{\ln x} - \frac{1}{x-1}\right) &= \lim_{x\to 1^+} \frac{\dfrac{d}{dx}[x-1-\ln x]}{\dfrac{d}{dx}[(x-1)\ln x]} \\
&= \lim_{x\to 1^+} \left[\frac{1-(1/x)}{(x-1)(1/x)+\ln x}\right] \\
&= \lim_{x\to 1^+} \left(\frac{x-1}{x-1+x\ln x}\right).
\end{aligned}$$

This limit also yields the indeterminate form $0/0$, so you can apply L'Hôpital's Rule again to obtain

$$\begin{aligned}
\lim_{x\to 1^+} \left(\frac{1}{\ln x} - \frac{1}{x-1}\right) &= \lim_{x\to 1^+} \left[\frac{1}{1+x(1/x)+\ln x}\right] \\
&= \frac{1}{2}.
\end{aligned}$$

Try It   Exploration A   Exploration B

The forms $0/0$, ∞/∞, $\infty - \infty$, $0 \cdot \infty$, 0^0, 1^∞, and ∞^0 have been identified as *indeterminate*. There are similar forms that you should recognize as "determinate."

$$\infty + \infty \to \infty \qquad \text{Limit is positive infinity.}$$

$$-\infty - \infty \to -\infty \qquad \text{Limit is negative infinity.}$$

$$0^\infty \to 0 \qquad \text{Limit is zero.}$$

$$0^{-\infty} \to \infty \qquad \text{Limit is positive infinity.}$$

(You are asked to verify two of these in Exercises 106 and 107.)

As a final comment, remember that L'Hôpital's Rule can be applied only to quotients leading to the indeterminate forms $0/0$ and ∞/∞. For instance, the following application of L'Hôpital's Rule is *incorrect*.

$$\lim_{x\to 0} \frac{e^x}{x} = \lim_{x\to 0} \frac{e^x}{1} = 1 \qquad \text{Incorrect use of L'Hôpital's Rule}$$

The reason this application is incorrect is that, even though the limit of the denominator is 0, the limit of the numerator is 1, which means that the hypotheses of L'Hôpital's Rule have not been satisfied.

Section 8.8 Improper Integrals

- **Evaluate an improper integral that has an infinite limit of integration.**
- **Evaluate an improper integral that has an infinite discontinuity.**

Improper Integrals with Infinite Limits of Integration

The definition of a definite integral

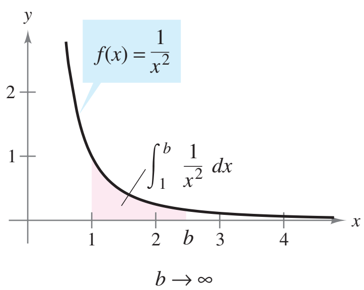
$$\int_a^b f(x)\,dx$$

requires that the interval $[a, b]$ be finite. Furthermore, the Fundamental Theorem of Calculus, by which you have been evaluating definite integrals, requires that f be continuous on $[a, b]$. In this section you will study a procedure for evaluating integrals that do not satisfy these requirements—usually because either one or both of the limits of integration are infinite, or f has a finite number of infinite discontinuities in the interval $[a, b]$. Integrals that possess either property are **improper integrals.** Note that a function f is said to have an **infinite discontinuity** at c if, *from the right or left,*

$$\lim_{x\to c} f(x) = \infty \quad \text{or} \quad \lim_{x\to c} f(x) = -\infty.$$

To get an idea of how to evaluate an improper integral, consider the integral

$$\int_1^b \frac{dx}{x^2} = -\frac{1}{x}\bigg]_1^b = -\frac{1}{b} + 1 = 1 - \frac{1}{b}$$



The unbounded region has an area of 1.
Figure 8.17

which can be interpreted as the area of the shaded region shown in Figure 8.17. Taking the limit as $b \to \infty$ produces

$$\int_1^\infty \frac{dx}{x^2} = \lim_{b\to\infty}\left(\int_1^b \frac{dx}{x^2}\right) = \lim_{b\to\infty}\left(1 - \frac{1}{b}\right) = 1.$$

This improper integral can be interpreted as the area of the *unbounded* region between the graph of $f(x) = 1/x^2$ and the x-axis (to the right of $x = 1$).

Definition of Improper Integrals with Infinite Integration Limits

1. If f is continuous on the interval $[a, \infty)$, then

$$\int_a^\infty f(x)\,dx = \lim_{b\to\infty}\int_a^b f(x)\,dx.$$

2. If f is continuous on the interval $(-\infty, b]$, then

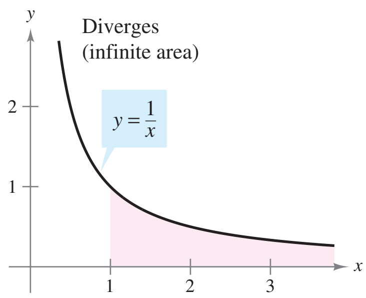
$$\int_{-\infty}^b f(x)\,dx = \lim_{a\to-\infty}\int_a^b f(x)\,dx.$$

3. If f is continuous on the interval $(-\infty, \infty)$, then

$$\int_{-\infty}^\infty f(x)\,dx = \int_{-\infty}^c f(x)\,dx + \int_c^\infty f(x)\,dx$$

where c is any real number (see Exercise 110).

In the first two cases, the improper integral **converges** if the limit exists—otherwise, the improper integral **diverges.** In the third case, the improper integral on the left diverges if either of the improper integrals on the right diverges.



This unbounded region has an infinite area.
Figure 8.18

Editable Graph

EXAMPLE 1 An Improper Integral That Diverges

Evaluate $\displaystyle\int_1^{\infty} \frac{dx}{x}$.

Solution

$$\begin{aligned}
\int_1^{\infty} \frac{dx}{x} &= \lim_{b\to\infty} \int_1^b \frac{dx}{x} && \text{Take limit as } b \to \infty. \\
&= \lim_{b\to\infty} \Big[\ln x\Big]_1^b && \text{Apply Log Rule.} \\
&= \lim_{b\to\infty} (\ln b - 0) && \text{Apply Fundamental Theorem of Calculus.} \\
&= \infty && \text{Evaluate limit.}
\end{aligned}$$

See Figure 8.18.

Try It Exploration A Exploration B Exploration C

NOTE Try comparing the regions shown in Figures 8.17 and 8.18. They look similar, yet the region in Figure 8.17 has a finite area of 1 and the region in Figure 8.18 has an infinite area.

EXAMPLE 2 Improper Integrals That Converge

Evaluate each improper integral.

a. $\displaystyle\int_0^{\infty} e^{-x}\,dx$ **b.** $\displaystyle\int_0^{\infty} \frac{1}{x^2+1}\,dx$

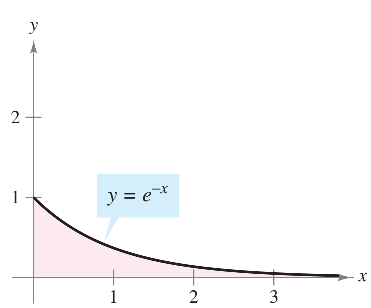
Solution

a. $$\begin{aligned}
\int_0^{\infty} e^{-x}\,dx &= \lim_{b\to\infty} \int_0^b e^{-x}\,dx \\
&= \lim_{b\to\infty} \Big[-e^{-x}\Big]_0^b \\
&= \lim_{b\to\infty} (-e^{-b} + 1) \\
&= 1
\end{aligned}$$

See Figure 8.19.

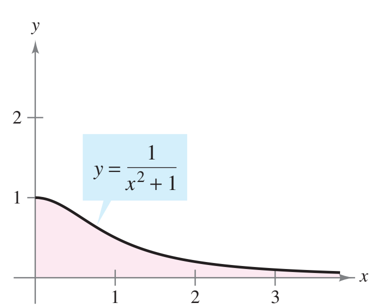
b. $$\begin{aligned}
\int_0^{\infty} \frac{1}{x^2+1}\,dx &= \lim_{b\to\infty} \int_0^b \frac{1}{x^2+1}\,dx \\
&= \lim_{b\to\infty} \Big[\arctan x\Big]_0^b \\
&= \lim_{b\to\infty} \arctan b \\
&= \frac{\pi}{2}
\end{aligned}$$

See Figure 8.20.



The area of the unbounded region is 1.
Figure 8.19

Editable Graph



The area of the unbounded region is $\pi/2$.
Figure 8.20

Editable Graph

Try It Exploration A

In the following example, note how L'Hôpital's Rule can be used to evaluate an improper integral.

EXAMPLE 3 Using L'Hôpital's Rule with an Improper Integral

Evaluate $\displaystyle\int_1^{\infty} (1 - x)e^{-x}\, dx$.

Solution Use integration by parts, with $dv = e^{-x}\, dx$ and $u = (1 - x)$.

$$\begin{aligned}\int (1 - x)e^{-x}\, dx &= -e^{-x}(1 - x) - \int e^{-x}\, dx \\ &= -e^{-x} + xe^{-x} + e^{-x} + C \\ &= xe^{-x} + C\end{aligned}$$

Now, apply the definition of an improper integral.

$$\begin{aligned}\int_1^{\infty} (1 - x)e^{-x}\, dx &= \lim_{b\to\infty} \Big[xe^{-x} \Big]_1^b \\ &= \left(\lim_{b\to\infty} \frac{b}{e^b}\right) - \frac{1}{e}\end{aligned}$$

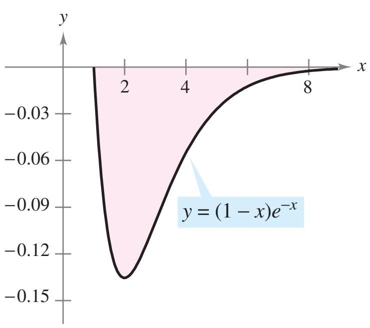
Finally, using L'Hôpital's Rule on the right-hand limit produces

$$\lim_{b\to\infty} \frac{b}{e^b} = \lim_{b\to\infty} \frac{1}{e^b} = 0$$

from which you can conclude that

$$\int_1^{\infty} (1 - x)e^{-x}\, dx = -\frac{1}{e}.$$

See Figure 8.21.



The area of the unbounded region is $|-1/e|$.
Figure 8.21

Editable Graph

Try It Exploration A

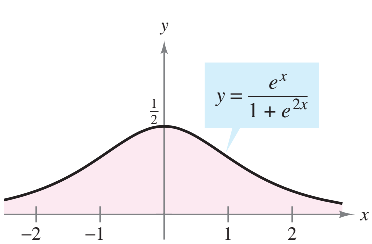
EXAMPLE 4 Infinite Upper and Lower Limits of Integration

Evaluate $\displaystyle\int_{-\infty}^{\infty} \frac{e^x}{1 + e^{2x}}\, dx$.

Solution Note that the integrand is continuous on $(-\infty, \infty)$. To evaluate the integral, you can break it into two parts, choosing $c = 0$ as a convenient value.

$$\begin{aligned}\int_{-\infty}^{\infty} \frac{e^x}{1 + e^{2x}}\, dx &= \int_{-\infty}^{0} \frac{e^x}{1 + e^{2x}}\, dx + \int_{0}^{\infty} \frac{e^x}{1 + e^{2x}}\, dx \\ &= \lim_{b\to-\infty} \Big[\arctan e^x \Big]_b^0 + \lim_{b\to\infty} \Big[\arctan e^x \Big]_0^b \\ &= \lim_{b\to-\infty} \left(\frac{\pi}{4} - \arctan e^b\right) + \lim_{b\to\infty} \left(\arctan e^b - \frac{\pi}{4}\right) \\ &= \frac{\pi}{4} - 0 + \frac{\pi}{2} - \frac{\pi}{4} \\ &= \frac{\pi}{2}\end{aligned}$$

See Figure 8.22.



The area of the unbounded region is $\pi/2$.
Figure 8.22

Editable Graph

Try It Exploration A

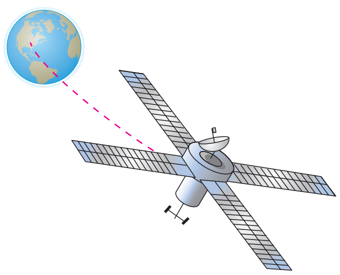
EXAMPLE 5 Sending a Space Module into Orbit

In Example 3 of Section 7.5, you found that it would require 10,000 mile-tons of work to propel a 15-metric-ton space module to a height of 800 miles above Earth. How much work is required to propel the module an unlimited distance away from Earth's surface?

Solution At first you might think that an infinite amount of work would be required. But if this were the case, it would be impossible to send rockets into outer space. Because this has been done, the work required must be finite. You can determine the work in the following manner. Using the integral of Example 3, Section 7.5, replace the upper bound of 4800 miles by ∞ and write

$$\begin{aligned} W &= \int_{4000}^{\infty} \frac{240{,}000{,}000}{x^2}\,dx \\ &= \lim_{b\to\infty}\left[-\frac{240{,}000{,}000}{x}\right]_{4000}^{b} \\ &= \lim_{b\to\infty}\left(-\frac{240{,}000{,}000}{b} + \frac{240{,}000{,}000}{4000}\right) \\ &= 60{,}000 \text{ mile-tons} \\ &\approx 6.984 \times 10^{11} \text{ foot-pounds.} \end{aligned}$$

See Figure 8.23.



The work required to move a space module an unlimited distance away from Earth is approximately 6.984×10^{11} foot-pounds.
Figure 8.23

Try It Exploration A

View the video to see the launching of the NASA SOLRAD-10 satellite.

Video

Improper Integrals with Infinite Discontinuities

The second basic type of improper integral is one that has an infinite discontinuity *at or between* the limits of integration.

Definition of Improper Integrals with Infinite Discontinuities

1. If f is continuous on the interval $[a, b)$ and has an infinite discontinuity at b, then

$$\int_a^b f(x)\,dx = \lim_{c\to b^-} \int_a^c f(x)\,dx.$$

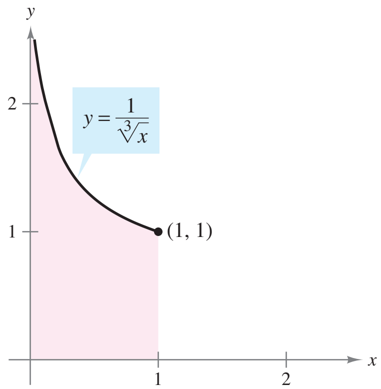
2. If f is continuous on the interval $(a, b]$ and has an infinite discontinuity at a, then

$$\int_a^b f(x)\,dx = \lim_{c\to a^+} \int_c^b f(x)\,dx.$$

3. If f is continuous on the interval $[a, b]$, except for some c in (a, b) at which f has an infinite discontinuity, then

$$\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx.$$

In the first two cases, the improper integral **converges** if the limit exists—otherwise, the improper integral **diverges.** In the third case, the improper integral on the left diverges if either of the improper integrals on the right diverges.



Infinite discontinuity at $x = 0$
Figure 8.24

Editable Graph

EXAMPLE 6 An Improper Integral with an Infinite Discontinuity

Evaluate $\displaystyle\int_0^1 \frac{dx}{\sqrt[3]{x}}$.

Solution The integrand has an infinite discontinuity at $x = 0$, as shown in Figure 8.24. You can evaluate this integral as shown below.

$$\begin{aligned}\int_0^1 x^{-1/3}\,dx &= \lim_{b\to 0^+}\left[\frac{x^{2/3}}{2/3}\right]_b^1\\ &= \lim_{b\to 0^+}\frac{3}{2}(1 - b^{2/3})\\ &= \frac{3}{2}\end{aligned}$$

Try It Exploration A Exploration B

EXAMPLE 7 An Improper Integral That Diverges

Evaluate $\displaystyle\int_0^2 \frac{dx}{x^3}$.

Solution Because the integrand has an infinite discontinuity at $x = 0$, you can write

$$\begin{aligned}\int_0^2 \frac{dx}{x^3} &= \lim_{b\to 0^+}\left[-\frac{1}{2x^2}\right]_b^2\\ &= \lim_{b\to 0^+}\left(-\frac{1}{8} + \frac{1}{2b^2}\right)\\ &= \infty.\end{aligned}$$

So, you can conclude that the improper integral diverges.

Try It Exploration A

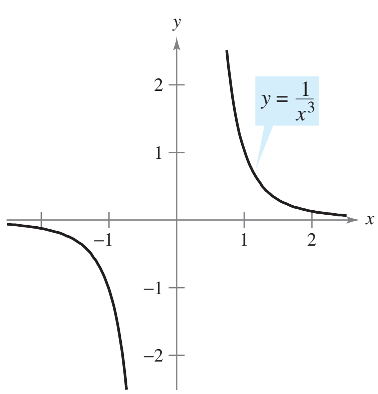
EXAMPLE 8 An Improper Integral with an Interior Discontinuity

Evaluate $\displaystyle\int_{-1}^2 \frac{dx}{x^3}$.

Solution This integral is improper because the integrand has an infinite discontinuity at the interior point $x = 0$, as shown in Figure 8.25. So, you can write

$$\int_{-1}^2 \frac{dx}{x^3} = \int_{-1}^0 \frac{dx}{x^3} + \int_0^2 \frac{dx}{x^3}.$$

From Example 7 you know that the second integral diverges. So, the original improper integral also diverges.



The improper integral $\displaystyle\int_{-1}^2 1/x^3\,dx$ diverges.
Figure 8.25

Editable Graph

Try It Exploration A

NOTE Remember to check for infinite discontinuities at interior points as well as endpoints when determining whether an integral is improper. For instance, if you had not recognized that the integral in Example 8 was improper, you would have obtained the *incorrect* result

$$\int_{-1}^2 \frac{dx}{x^3} = \left[\frac{-1}{2x^2}\right]_{-1}^2 = -\frac{1}{8} + \frac{1}{2} = \frac{3}{8}.$$

Incorrect evaluation

The integral in the next example is improper for *two* reasons. One limit of integration is infinite, and the integrand has an infinite discontinuity at the outer limit of integration.

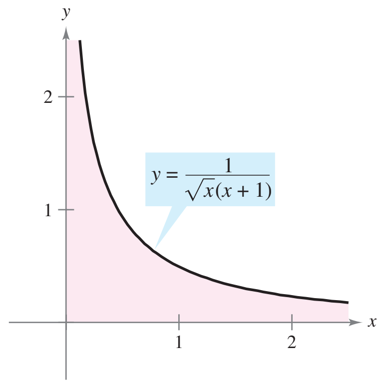
EXAMPLE 9 A Doubly Improper Integral

Evaluate $\displaystyle\int_0^{\infty} \frac{dx}{\sqrt{x}(x+1)}$.

Solution To evaluate this integral, split it at a convenient point (say, $x = 1$) and write

$$\begin{aligned}\int_0^{\infty} \frac{dx}{\sqrt{x}(x+1)} &= \int_0^1 \frac{dx}{\sqrt{x}(x+1)} + \int_1^{\infty} \frac{dx}{\sqrt{x}(x+1)}\\ &= \lim_{b\to 0^+}\Big[2\arctan\sqrt{x}\Big]_b^1 + \lim_{c\to\infty}\Big[2\arctan\sqrt{x}\Big]_1^c\\ &= 2\left(\frac{\pi}{4}\right) - 0 + 2\left(\frac{\pi}{2}\right) - 2\left(\frac{\pi}{4}\right)\\ &= \pi.\end{aligned}$$

See Figure 8.26.



The area of the unbounded region is π.
Figure 8.26

Editable Graph

Try It | Exploration A | Exploration B | Open Exploration

EXAMPLE 10 An Application Involving Arc Length

Use the formula for arc length to show that the circumference of the circle $x^2 + y^2 = 1$ is 2π.

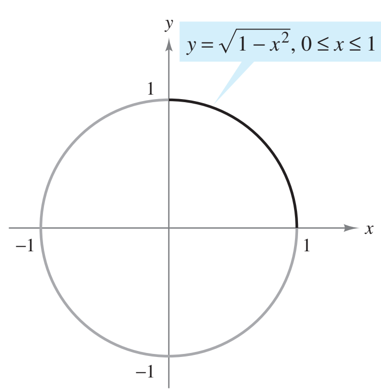
Solution To simplify the work, consider the quarter circle given by $y = \sqrt{1 - x^2}$, where $0 \le x \le 1$. The function y is differentiable for any x in this interval except $x = 1$. Therefore, the arc length of the quarter circle is given by the improper integral

$$\begin{aligned}s &= \int_0^1 \sqrt{1 + (y')^2}\,dx\\ &= \int_0^1 \sqrt{1 + \left(\frac{-x}{\sqrt{1-x^2}}\right)^2}\,dx\\ &= \int_0^1 \frac{dx}{\sqrt{1-x^2}}.\end{aligned}$$

This integral is improper because it has an infinite discontinuity at $x = 1$. So, you can write

$$\begin{aligned}s &= \int_0^1 \frac{dx}{\sqrt{1-x^2}}\\ &= \lim_{b\to 1^-}\Big[\arcsin x\Big]_0^b\\ &= \frac{\pi}{2} - 0\\ &= \frac{\pi}{2}.\end{aligned}$$

Finally, multiplying by 4, you can conclude that the circumference of the circle is $4s = 2\pi$, as shown in Figure 8.27.



The circumference of the circle is 2π.
Figure 8.27

Editable Graph

Try It | Exploration A

This section concludes with a useful theorem describing the convergence or divergence of a common type of improper integral. The proof of this theorem is left as an exercise (see Exercise 49).

THEOREM 8.5 A Special Type of Improper Integral

$$\int_1^{\infty} \frac{dx}{x^p} = \begin{cases} \dfrac{1}{p-1}, & \text{if } p > 1 \\ \text{diverges}, & \text{if } p \le 1 \end{cases}$$

EXAMPLE 11 An Application Involving A Solid of Revolution

FOR FURTHER INFORMATION For further investigation of solids that have finite volumes and infinite surface areas, see the article "Supersolids: Solids Having Finite Volume and Infinite Surfaces" by William P. Love in *Mathematics Teacher*.

MathArticle

The solid formed by revolving (about the x-axis) the *unbounded* region lying between the graph of $f(x) = 1/x$ and the x-axis ($x \ge 1$) is called **Gabriel's Horn.** (See Figure 8.28.) Show that this solid has a finite volume and an infinite surface area.

Solution Using the disk method and Theorem 8.5, you can determine the volume to be

$$V = \pi \int_1^{\infty} \left(\frac{1}{x}\right)^2 dx \qquad \text{Theorem 8.5, } p = 2 > 1$$

$$= \pi\left(\frac{1}{2-1}\right) = \pi.$$

The surface area is given by

$$S = 2\pi \int_1^{\infty} f(x)\sqrt{1 + [f'(x)]^2}\, dx = 2\pi \int_1^{\infty} \frac{1}{x}\sqrt{1 + \frac{1}{x^4}}\, dx.$$

Because

$$\sqrt{1 + \frac{1}{x^4}} > 1$$

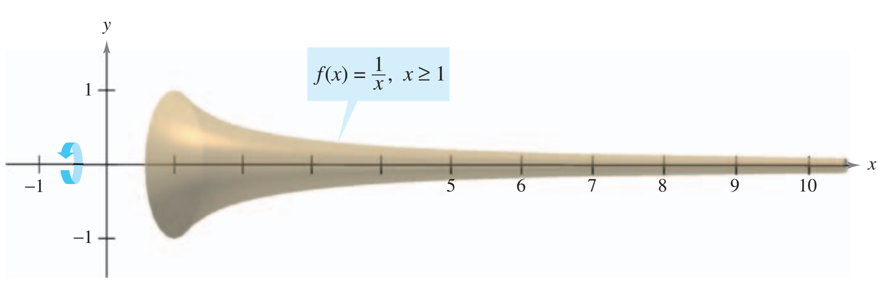
on the interval $[1, \infty)$, and the improper integral

$$\int_1^{\infty} \frac{1}{x}\, dx$$

diverges, you can conclude that the improper integral

$$\int_1^{\infty} \frac{1}{x}\sqrt{1 + \frac{1}{x^4}}\, dx$$

also diverges. (See Exercise 52.) So, the surface area is infinite.



Gabriel's Horn has a finite volume and an infinite surface area.
Figure 8.28

FOR FURTHER INFORMATION To learn about another function that has a finite volume and an infinite surface area, see the article "Gabriel's Wedding Cake" by Julian F. Fleron in *The College Mathematics Journal*.

MathArticle

Rotatable Graph

Try It

Exploration A

Section 9.1 Sequences

- **List the terms of a sequence.**
- **Determine whether a sequence converges or diverges.**
- **Write a formula for the nth term of a sequence.**
- **Use properties of monotonic sequences and bounded sequences.**

Sequences

In mathematics, the word "sequence" is used in much the same way as in ordinary English. To say that a collection of objects or events is *in sequence* usually means that the collection is ordered so that it has an identified first member, second member, third member, and so on.

Mathematically, a **sequence** is defined as a function whose domain is the set of positive integers. Although a sequence is a function, it is common to represent sequences by subscript notation rather than by the standard function notation. For instance, in the sequence

$$\begin{array}{ccccccc} 1, & 2, & 3, & 4, & \ldots, & n, & \ldots \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ a_1, & a_2, & a_3, & a_4, & \ldots, & a_n, & \ldots \end{array} \qquad \text{Sequence}$$

1 is mapped onto a_1, 2 is mapped onto a_2, and so on. The numbers $a_1, a_2, a_3, \ldots, a_n, \ldots$ are the **terms** of the sequence. The number a_n is the ***n*th term** of the sequence, and the entire sequence is denoted by $\{a_n\}$.

EXPLORATION

Finding Patterns Describe a pattern for each of the following sequences. Then use your description to write a formula for the nth term of each sequence. As n increases, do the terms appear to be approaching a limit? Explain your reasoning.

a. $1, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \ldots$

b. $1, \frac{1}{2}, \frac{1}{6}, \frac{1}{24}, \frac{1}{120}, \ldots$

c. $10, \frac{10}{3}, \frac{10}{6}, \frac{10}{10}, \frac{10}{15}, \ldots$

d. $\frac{1}{4}, \frac{4}{9}, \frac{9}{16}, \frac{16}{25}, \frac{25}{36}, \ldots$

e. $\frac{3}{7}, \frac{5}{10}, \frac{7}{13}, \frac{9}{16}, \frac{11}{19}, \ldots$

NOTE Occasionally, it is convenient to begin a sequence with a_0, so that the terms of the sequence become $a_0, a_1, a_2, a_3, \ldots, a_n, \ldots$

EXAMPLE 1 Listing the Terms of a Sequence

a. The terms of the sequence $\{a_n\} = \{3 + (-1)^n\}$ are

$$\begin{array}{ccccc} 3 + (-1)^1, & 3 + (-1)^2, & 3 + (-1)^3, & 3 + (-1)^4, & \ldots \\ 2, & 4, & 2, & 4, & \ldots. \end{array}$$

b. The terms of the sequence $\{b_n\} = \left\{\dfrac{n}{1-2n}\right\}$ are

$$\begin{array}{ccccc} \dfrac{1}{1-2\cdot 1}, & \dfrac{2}{1-2\cdot 2}, & \dfrac{3}{1-2\cdot 3}, & \dfrac{4}{1-2\cdot 4}, & \ldots \\ -1, & -\dfrac{2}{3}, & -\dfrac{3}{5}, & -\dfrac{4}{7}, & \ldots. \end{array}$$

c. The terms of the sequence $\{c_n\} = \left\{\dfrac{n^2}{2^n-1}\right\}$ are

$$\begin{array}{ccccc} \dfrac{1^2}{2^1-1}, & \dfrac{2^2}{2^2-1}, & \dfrac{3^2}{2^3-1}, & \dfrac{4^2}{2^4-1}, & \ldots \\ \dfrac{1}{1}, & \dfrac{4}{3}, & \dfrac{9}{7}, & \dfrac{16}{15}, & \ldots. \end{array}$$

d. The terms of the **recursively defined** sequence $\{d_n\}$, where $d_1 = 25$ and $d_{n+1} = d_n - 5$ are

$$25, \quad 25 - 5 = 20, \quad 20 - 5 = 15, \quad 15 - 5 = 10, \ldots.$$

STUDY TIP Some sequences are defined recursively. To define a sequence recursively, you need to be given one or more of the first few terms. All other terms of the sequence are then defined using previous terms, as shown in Example 1(d).

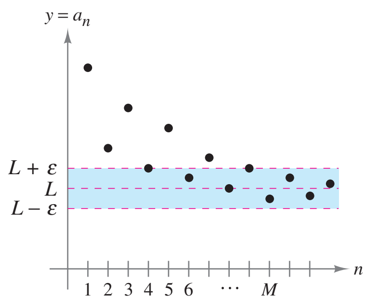
Try It Exploration A

Limit of a Sequence

The primary focus of this chapter concerns sequences whose terms approach limiting values. Such sequences are said to **converge.** For instance, the sequence $\{1/2^n\}$

$$\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \frac{1}{32}, \ldots$$

converges to 0, as indicated in the following definition.



For $n > M$, the terms of the sequence all lie within ε units of L.
Figure 9.1

Definition of the Limit of a Sequence

Let L be a real number. The **limit** of a sequence $\{a_n\}$ is L, written as

$$\lim_{n\to\infty} a_n = L$$

if for each $\varepsilon > 0$, there exists $M > 0$ such that $|a_n - L| < \varepsilon$ whenever $n > M$. If the limit L of a sequence exists, then the sequence **converges** to L. If the limit of a sequence does not exist, then the sequence **diverges.**

Graphically, this definition says that eventually (for $n > M$ and $\varepsilon > 0$) the terms of a sequence that converges to L will lie within the band between the lines $y = L + \varepsilon$ and $y = L - \varepsilon$, as shown in Figure 9.1.

If a sequence $\{a_n\}$ agrees with a function f at every positive integer, and if $f(x)$ approaches a limit L as $x \to \infty$, the sequence must converge to the same limit L.

THEOREM 9.1 Limit of a Sequence

Let L be a real number. Let f be a function of a real variable such that

$$\lim_{x\to\infty} f(x) = L.$$

If $\{a_n\}$ is a sequence such that $f(n) = a_n$ for every positive integer n, then

$$\lim_{n\to\infty} a_n = L.$$

NOTE There are different ways in which a sequence can fail to have a limit. One way is that the terms of the sequence increase without bound or decrease without bound. These cases are written symbolically as follows.

Terms increase without bound:

$$\lim_{n\to\infty} a_n = \infty$$

Terms decrease without bound:

$$\lim_{n\to\infty} a_n = -\infty$$

EXAMPLE 2 Finding the Limit of a Sequence

Find the limit of the sequence whose nth term is

$$a_n = \left(1 + \frac{1}{n}\right)^n.$$

Solution In Theorem 5.15, you learned that

$$\lim_{x\to\infty} \left(1 + \frac{1}{x}\right)^x = e.$$

So, you can apply Theorem 9.1 to conclude that

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n = e.$$

Try It | Exploration A | Exploration B

The following properties of limits of sequences parallel those given for limits of functions of a real variable in Section 1.3.

THEOREM 9.2 Properties of Limits of Sequences

Let $\lim_{n\to\infty} a_n = L$ and $\lim_{n\to\infty} b_n = K$.

1. $\lim_{n\to\infty} (a_n \pm b_n) = L \pm K$

2. $\lim_{n\to\infty} ca_n = cL$, c is any real number

3. $\lim_{n\to\infty} (a_n b_n) = LK$

4. $\lim_{n\to\infty} \dfrac{a_n}{b_n} = \dfrac{L}{K}$, $b_n \neq 0$ and $K \neq 0$

EXAMPLE 3 Determining Convergence or Divergence

a. Because the sequence $\{a_n\} = \{3 + (-1)^n\}$ has terms

$$2, 4, 2, 4, \ldots$$

See Example 1(a), page 594.

that alternate between 2 and 4, the limit

$$\lim_{n\to\infty} a_n$$

does not exist. So, the sequence diverges.

b. For $\{b_n\} = \left\{\dfrac{n}{1-2n}\right\}$, divide the numerator and denominator by n to obtain

$$\lim_{n\to\infty} \frac{n}{1-2n} = \lim_{n\to\infty} \frac{1}{(1/n) - 2} = -\frac{1}{2}$$

See Example 1(b), page 594.

which implies that the sequence converges to $-\frac{1}{2}$.

Try It | Exploration A | Open Exploration | Technology

EXAMPLE 4 Using L'Hôpital's Rule to Determine Convergence

Show that the sequence whose nth term is $a_n = \dfrac{n^2}{2^n - 1}$ converges.

Solution Consider the function of a real variable

$$f(x) = \frac{x^2}{2^x - 1}.$$

Applying L'Hôpital's Rule twice produces

$$\lim_{x\to\infty} \frac{x^2}{2^x - 1} = \lim_{x\to\infty} \frac{2x}{(\ln 2)2^x} = \lim_{x\to\infty} \frac{2}{(\ln 2)^2 2^x} = 0.$$

Because $f(n) = a_n$ for every positive integer, you can apply Theorem 9.1 to conclude that

$$\lim_{n\to\infty} \frac{n^2}{2^n - 1} = 0.$$

See Example 1(c), page 594.

So, the sequence converges to 0.

Try It | Exploration A | Exploration B

TECHNOLOGY Use a graphing utility to graph the function in Example 4. Notice that as x approaches infinity, the value of the function gets closer and closer to 0. If you have access to a graphing utility that can generate terms of a sequence, try using it to calculate the first 20 terms of the sequence in Example 4. Then view the terms to observe numerically that the sequence converges to 0.

The symbol $n!$ (read "n factorial") is used to simplify some of the formulas developed in this chapter. Let n be a positive integer; then ***n* factorial** is defined as

$$n! = 1 \cdot 2 \cdot 3 \cdot 4 \cdot \cdot \cdot (n - 1) \cdot n.$$

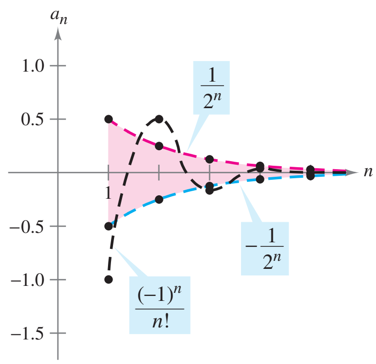
As a special case, **zero factorial** is defined as $0! = 1$. From this definition, you can see that $1! = 1$, $2! = 1 \cdot 2 = 2$, $3! = 1 \cdot 2 \cdot 3 = 6$, and so on. Factorials follow the same conventions for order of operations as exponents. That is, just as $2x^3$ and $(2x)^3$ imply different orders of operations, $2n!$ and $(2n)!$ imply the following orders.

$$2n! = 2(n!) = 2(1 \cdot 2 \cdot 3 \cdot 4 \cdot \cdot \cdot n)$$

and

$$(2n)! = 1 \cdot 2 \cdot 3 \cdot 4 \cdot \cdot \cdot n \cdot (n + 1) \cdot \cdot \cdot 2n$$

Another useful limit theorem that can be rewritten for sequences is the Squeeze Theorem from Section 1.3.



For $n \geq 4$, $(-1)^n/n!$ is squeezed between $-1/2^n$ and $1/2^n$.
Figure 9.2

THEOREM 9.3 Squeeze Theorem for Sequences

If

$$\lim_{n\to\infty} a_n = L = \lim_{n\to\infty} b_n$$

and there exists an integer N such that $a_n \leq c_n \leq b_n$ for all $n > N$, then

$$\lim_{n\to\infty} c_n = L.$$

EXAMPLE 5 Using the Squeeze Theorem

Show that the sequence $\{c_n\} = \left\{(-1)^n \frac{1}{n!}\right\}$ converges, and find its limit.

Solution To apply the Squeeze Theorem, you must find two convergent sequences that can be related to the given sequence. Two possibilities are $a_n = -1/2^n$ and $b_n = 1/2^n$, both of which converge to 0. By comparing the term $n!$ with 2^n, you can see that

$$n! = 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6 \cdot \cdot \cdot n = 24 \cdot \underbrace{5 \cdot 6 \cdot \cdot \cdot n}_{n\,-\,4 \text{ factors}} \qquad (n \geq 4)$$

and

$$2^n = 2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 \cdot \cdot \cdot 2 = 16 \cdot \underbrace{2 \cdot 2 \cdot \cdot \cdot 2}_{n\,-\,4 \text{ factors}}. \qquad (n \geq 4)$$

This implies that for $n \geq 4$, $2^n < n!$, and you have

$$\frac{-1}{2^n} \leq (-1)^n \frac{1}{n!} \leq \frac{1}{2^n}, \quad n \geq 4$$

as shown in Figure 9.2. So, by the Squeeze Theorem it follows that

$$\lim_{n\to\infty} (-1)^n \frac{1}{n!} = 0.$$

Try It | Exploration A | Exploration B

NOTE Example 5 suggests something about the rate at which $n!$ increases as $n \to \infty$. As Figure 9.2 suggests, both $1/2^n$ and $1/n!$ approach 0 as $n \to \infty$. Yet $1/n!$ approaches 0 so much faster than $1/2^n$ does that

$$\lim_{n\to\infty} \frac{1/n!}{1/2^n} = \lim_{n\to\infty} \frac{2^n}{n!} = 0.$$

In fact, it can be shown that for any fixed number k,

$$\lim_{n\to\infty} \frac{k^n}{n!} = 0.$$

This means that *the factorial function grows faster than any exponential function.*

In Example 5, the sequence $\{c_n\}$ has both positive and negative terms. For this sequence, it happens that the sequence of absolute values, $\{|c_n|\}$, also converges to 0. You can show this by the Squeeze Theorem using the inequality

$$0 \le \frac{1}{n!} \le \frac{1}{2^n}, \quad n \ge 4.$$

In such cases, it is often convenient to consider the sequence of absolute values—and then apply Theorem 9.4, which states that if the absolute value sequence converges to 0, the original signed sequence also converges to 0.

THEOREM 9.4 Absolute Value Theorem

For the sequence $\{a_n\}$, if

$$\lim_{n\to\infty} |a_n| = 0 \qquad \text{then} \qquad \lim_{n\to\infty} a_n = 0.$$

Proof Consider the two sequences $\{|a_n|\}$ and $\{-|a_n|\}$. Because both of these sequences converge to 0 and

$$-|a_n| \le a_n \le |a_n|$$

you can use the Squeeze Theorem to conclude that $\{a_n\}$ converges to 0.

Pattern Recognition for Sequences

Sometimes the terms of a sequence are generated by some rule that does not explicitly identify the *n*th term of the sequence. In such cases, you may be required to discover a *pattern* in the sequence and to describe the *n*th term. Once the *n*th term has been specified, you can investigate the convergence or divergence of the sequence.

EXAMPLE 6 Finding the *n*th Term of a Sequence

Find a sequence $\{a_n\}$ whose first five terms are

$$\frac{2}{1}, \frac{4}{3}, \frac{8}{5}, \frac{16}{7}, \frac{32}{9}, \ldots$$

and then determine whether the particular sequence you have chosen converges or diverges.

Solution First, note that the numerators are successive powers of 2, and the denominators form the sequence of positive odd integers. By comparing a_n with n, you have the following pattern.

$$\frac{2^1}{1}, \frac{2^2}{3}, \frac{2^3}{5}, \frac{2^4}{7}, \frac{2^5}{9}, \ldots, \frac{2^n}{2n-1}$$

Using L'Hôpital's Rule to evaluate the limit of $f(x) = 2^x/(2x - 1)$, you obtain

$$\lim_{x\to\infty} \frac{2^x}{2x-1} = \lim_{x\to\infty} \frac{2^x(\ln 2)}{2} = \infty \quad \Rightarrow \quad \lim_{n\to\infty} \frac{2^n}{2n-1} = \infty.$$

So, the sequence diverges.

Try It Exploration A

Without a specific rule for generating the terms of a sequence or some knowledge of the context in which the terms of the sequence are obtained, it is not possible to determine the convergence or divergence of the sequence merely from its first several terms. For instance, although the first three terms of the following four sequences are identical, the first two sequences converge to 0, the third sequence converges to $\frac{1}{9}$, and the fourth sequence diverges.

$$\{a_n\}: \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \ldots, \frac{1}{2^n}, \ldots$$

$$\{b_n\}: \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{15}, \ldots, \frac{6}{(n+1)(n^2-n+6)}, \ldots$$

$$\{c_n\}: \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{7}{62}, \ldots, \frac{n^2-3n+3}{9n^2-25n+18}, \ldots$$

$$\{d_n\}: \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, 0, \ldots, \frac{-n(n+1)(n-4)}{6(n^2+3n-2)}, \ldots$$

The process of determining an nth term from the pattern observed in the first several terms of a sequence is an example of *inductive reasoning*.

EXAMPLE 7 Finding the nth Term of a Sequence

Determine an nth term for a sequence whose first five terms are

$$-\frac{2}{1}, \frac{8}{2}, -\frac{26}{6}, \frac{80}{24}, -\frac{242}{120}, \ldots$$

and then decide whether the sequence converges or diverges.

Solution Note that the numerators are 1 less than 3^n. So, you can reason that the numerators are given by the rule $3^n - 1$. Factoring the denominators produces

$$\begin{aligned} 1 &= 1 \\ 2 &= 1 \cdot 2 \\ 6 &= 1 \cdot 2 \cdot 3 \\ 24 &= 1 \cdot 2 \cdot 3 \cdot 4 \\ 120 &= 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot \cdot \cdot . \end{aligned}$$

This suggests that the denominators are represented by $n!$. Finally, because the signs alternate, you can write the nth term as

$$a_n = (-1)^n \left(\frac{3^n - 1}{n!}\right).$$

From the discussion about the growth of $n!$, it follows that

$$\lim_{n\to\infty} |a_n| = \lim_{n\to\infty} \frac{3^n - 1}{n!} = 0.$$

Applying Theorem 9.4, you can conclude that

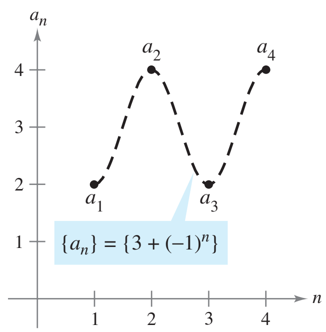
$$\lim_{n\to\infty} a_n = 0.$$

So, the sequence $\{a_n\}$ converges to 0.

Try It   Exploration A

Monotonic Sequences and Bounded Sequences

So far you have determined the convergence of a sequence by finding its limit. Even if you cannot determine the limit of a particular sequence, it still may be useful to know whether the sequence converges. Theorem 9.5 provides a test for convergence of sequences without determining the limit. First, some preliminary definitions are given.



(a) Not monotonic

Definition of a Monotonic Sequence

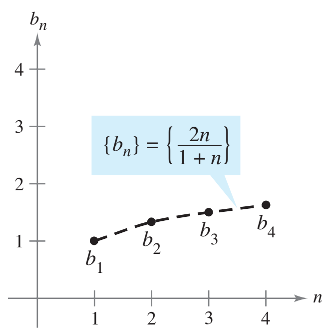
A sequence $\{a_n\}$ is **monotonic** if its terms are nondecreasing

$$a_1 \le a_2 \le a_3 \le \cdots \le a_n \le \cdots$$

or if its terms are nonincreasing

$$a_1 \ge a_2 \ge a_3 \ge \cdots \ge a_n \ge \cdots.$$

EXAMPLE 8 Determining Whether a Sequence Is Monotonic



(b) Monotonic

Determine whether each sequence having the given nth term is monotonic.

a. $a_n = 3 + (-1)^n$ **b.** $b_n = \dfrac{2n}{1+n}$ **c.** $c_n = \dfrac{n^2}{2^n - 1}$

Solution

a. This sequence alternates between 2 and 4. So, it is not monotonic.

b. This sequence is monotonic because each successive term is larger than its predecessor. To see this, compare the terms b_n and b_{n+1}. [Note that, because n is positive, you can multiply each side of the inequality by $(1 + n)$ and $(2 + n)$ without reversing the inequality sign.]

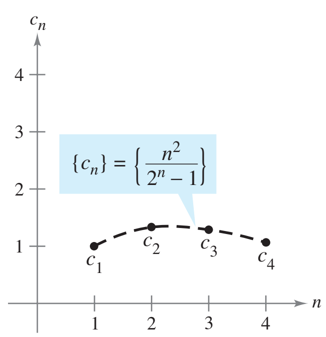
$$b_n = \frac{2n}{1+n} \overset{?}{<} \frac{2(n+1)}{1+(n+1)} = b_{n+1}$$

$$2n(2+n) \overset{?}{<} (1+n)(2n+2)$$

$$4n + 2n^2 \overset{?}{<} 2 + 4n + 2n^2$$

$$0 < 2$$

Starting with the final inequality, which is valid, you can reverse the steps to conclude that the original inequality is also valid.



(c) Not monotonic

Figure 9.3

c. This sequence is not monotonic, because the second term is larger than the first term, and larger than the third. (Note that if you drop the first term, the remaining sequence $c_2, c_3, c_4, \ldots$ is monotonic.)

Figure 9.3 graphically illustrates these three sequences.

Try It Exploration A Exploration B

NOTE In Example 8(b), another way to see that the sequence is monotonic is to argue that the derivative of the corresponding differentiable function $f(x) = 2x/(1 + x)$ is positive for all x. This implies that f is increasing, which in turn implies that $\{a_n\}$ is increasing.

NOTE All three sequences shown in Figure 9.3 are bounded. To see this, consider the following.

$$2 \le a_n \le 4$$
$$1 \le b_n \le 2$$
$$0 \le c_n \le \frac{4}{3}$$

Definition of a Bounded Sequence

1. A sequence $\{a_n\}$ is **bounded above** if there is a real number M such that $a_n \le M$ for all n. The number M is called an **upper bound** of the sequence.
2. A sequence $\{a_n\}$ is **bounded below** if there is a real number N such that $N \le a_n$ for all n. The number N is called a **lower bound** of the sequence.
3. A sequence $\{a_n\}$ is **bounded** if it is bounded above and bounded below.

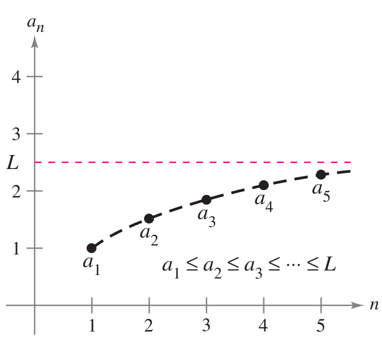
One important property of the real numbers is that they are **complete.** Informally, this means that there are no holes or gaps on the real number line. (The set of rational numbers does not have the completeness property.) The completeness axiom for real numbers can be used to conclude that if a sequence has an upper bound, it must have a **least upper bound** (an upper bound that is smaller than all other upper bounds for the sequence). For example, the least upper bound of the sequence $\{a_n\} = \{n/(n+1)\}$,

$$\frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \frac{4}{5}, \ldots, \frac{n}{n+1}, \ldots$$

is 1. The completeness axiom is used in the proof of Theorem 9.5.

THEOREM 9.5 Bounded Monotonic Sequences

If a sequence $\{a_n\}$ is bounded and monotonic, then it converges.



Every bounded nondecreasing sequence converges.
Figure 9.4

Proof Assume that the sequence is nondecreasing, as shown in Figure 9.4. For the sake of simplicity, also assume that each term in the sequence is positive. Because the sequence is bounded, there must exist an upper bound M such that

$$a_1 \le a_2 \le a_3 \le \cdots \le a_n \le \cdots \le M.$$

From the completeness axiom, it follows that there is a least upper bound L such that

$$a_1 \le a_2 \le a_3 \le \cdots \le a_n \le \cdots \le L.$$

For $\varepsilon > 0$, it follows that $L - \varepsilon < L$, and therefore $L - \varepsilon$ cannot be an upper bound for the sequence. Consequently, at least one term of $\{a_n\}$ is greater than $L - \varepsilon$. That is, $L - \varepsilon < a_N$ for some positive integer N. Because the terms of $\{a_n\}$ are nondecreasing, it follows that $a_N \le a_n$ for $n > N$. You now know that $L - \varepsilon < a_N \le a_n \le L < L + \varepsilon$, for every $n > N$. It follows that $|a_n - L| < \varepsilon$ for $n > N$, which by definition means that $\{a_n\}$ converges to L. The proof for a nonincreasing sequence is similar.

EXAMPLE 9 Bounded and Monotonic Sequences

a. The sequence $\{a_n\} = \{1/n\}$ is both bounded and monotonic and so, by Theorem 9.5, must converge.

b. The divergent sequence $\{b_n\} = \{n^2/(n+1)\}$ is monotonic, but not bounded. (It is bounded below.)

c. The divergent sequence $\{c_n\} = \{(-1)^n\}$ is bounded, but not monotonic.

Try It

Exploration A Exploration B Exploration C Exploration D

Section 9.2 Series and Convergence

- **Understand the definition of a convergent infinite series.**
- **Use properties of infinite geometric series.**
- **Use the *n*th-Term Test for Divergence of an infinite series.**

Infinite Series

One important application of infinite sequences is in representing "infinite summations." Informally, if $\{a_n\}$ is an infinite sequence, then

$$\sum_{n=1}^{\infty} a_n = a_1 + a_2 + a_3 + \cdot \cdot \cdot + a_n + \cdot \cdot \cdot$$

Infinite series

is an **infinite series** (or simply a **series**). The numbers a_1, a_2, a_3, are the **terms** of the series. For some series it is convenient to begin the index at $n = 0$ (or some other integer). As a typesetting convention, it is common to represent an infinite series as simply $\Sigma\, a_n$. In such cases, the starting value for the index must be taken from the context of the statement.

To find the sum of an infinite series, consider the following **sequence of partial sums.**

$$\begin{aligned} S_1 &= a_1 \\ S_2 &= a_1 + a_2 \\ S_3 &= a_1 + a_2 + a_3 \\ &\vdots \\ S_n &= a_1 + a_2 + a_3 + \cdot \cdot \cdot + a_n \end{aligned}$$

If this sequence of partial sums converges, the series is said to converge and has the sum indicated in the following definition.

Infinite Series

The study of infinite series was considered a novelty in the fourteenth century. Logician Richard Suiseth, whose nickname was Calculator, solved this problem.

If throughout the first half of a given time interval a variation continues at a certain intensity, throughout the next quarter of the interval at double the intensity, throughout the following eighth at triple the intensity and so ad infinitum; then the average intensity for the whole interval will be the intensity of the variation during the second subinterval (or double the intensity).

This is the same as saying that the sum of the infinite series

$$\frac{1}{2} + \frac{2}{4} + \frac{3}{8} + \cdot \cdot \cdot + \frac{n}{2^n} + \cdot \cdot \cdot$$

is 2.

Definitions of Convergent and Divergent Series

For the infinite series $\sum_{n=1}^{\infty} a_n$, the ***n*th partial sum** is given by

$$S_n = a_1 + a_2 + \cdot \cdot \cdot + a_n.$$

If the sequence of partial sums $\{S_n\}$ converges to S, then the series $\sum_{n=1}^{\infty} a_n$ **converges.** The limit S is called the **sum of the series.**

$$S = a_1 + a_2 + \cdot \cdot \cdot + a_n + \cdot \cdot \cdot$$

If $\{S_n\}$ diverges, then the series **diverges.**

STUDY TIP As you study this chapter, you will see that there are two basic questions involving infinite series. Does a series converge or does it diverge? If a series converges, what is its sum? These questions are not always easy to answer, especially the second one.

EXPLORATION

Finding the Sum of an Infinite Series Find the sum of each infinite series. Explain your reasoning.

a. $0.1 + 0.01 + 0.001 + 0.0001 + \cdot \cdot \cdot$

b. $\frac{3}{10} + \frac{3}{100} + \frac{3}{1000} + \frac{3}{10{,}000} + \cdot \cdot \cdot$

c. $1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \cdot \cdot \cdot$

d. $\frac{15}{100} + \frac{15}{10{,}000} + \frac{15}{1{,}000{,}000} + \cdot \cdot \cdot$

TECHNOLOGY Figure 9.5 shows the first 15 partial sums of the infinite series in Example 1(a). Notice how the values appear to approach the line $y = 1$.

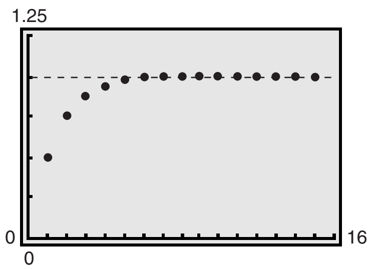


Figure 9.5

NOTE You can geometrically determine the partial sums of the series in Example 1(a) using Figure 9.6.

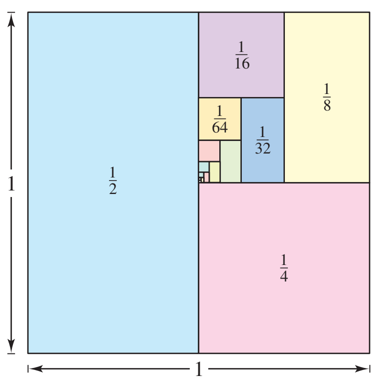


Figure 9.6

EXAMPLE 1 Convergent and Divergent Series

a. The series

$$\sum_{n=1}^{\infty} \frac{1}{2^n} = \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \cdot \cdot \cdot$$

has the following partial sums.

$$S_1 = \frac{1}{2}$$

$$S_2 = \frac{1}{2} + \frac{1}{4} = \frac{3}{4}$$

$$S_3 = \frac{1}{2} + \frac{1}{4} + \frac{1}{8} = \frac{7}{8}$$

$$\vdots$$

$$S_n = \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \cdot \cdot \cdot + \frac{1}{2^n} = \frac{2^n - 1}{2^n}$$

Because

$$\lim_{n\to\infty} \frac{2^n - 1}{2^n} = 1$$

it follows that the series converges and its sum is 1.

b. The nth partial sum of the series

$$\sum_{n=1}^{\infty} \left(\frac{1}{n} - \frac{1}{n+1}\right) = \left(1 - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) + \left(\frac{1}{3} - \frac{1}{4}\right) + \cdot \cdot \cdot$$

is given by

$$S_n = 1 - \frac{1}{n+1}.$$

Because the limit of S_n is 1, the series converges and its sum is 1.

c. The series

$$\sum_{n=1}^{\infty} 1 = 1 + 1 + 1 + 1 + \cdot \cdot \cdot$$

diverges because $S_n = n$ and the sequence of partial sums diverges.

Try It | Exploration A | Exploration B | Exploration C

Technology

The series in Example 1(b) is a **telescoping series** of the form

$$(b_1 - b_2) + (b_2 - b_3) + (b_3 - b_4) + (b_4 - b_5) + \cdot \cdot \cdot. \qquad \text{Telescoping series}$$

Note that b_2 is canceled by the second term, b_3 is canceled by the third term, and so on. Because the nth partial sum of this series is

$$S_n = b_1 - b_{n+1}$$

it follows that a telescoping series will converge if and only if b_n approaches a finite number as $n \to \infty$. Moreover, if the series converges, its sum is

$$S = b_1 - \lim_{n\to\infty} b_{n+1}.$$

FOR FURTHER INFORMATION To learn more about the partial sums of infinite series, see the article "Six Ways to Sum a Series" by Dan Kalman in *The College Mathematics Journal*.

MathArticle

EXAMPLE 2 Writing a Series in Telescoping Form

Find the sum of the series $\displaystyle\sum_{n=1}^{\infty} \frac{2}{4n^2 - 1}$.

Solution
Using partial fractions, you can write

$$a_n = \frac{2}{4n^2 - 1} = \frac{2}{(2n - 1)(2n + 1)} = \frac{1}{2n - 1} - \frac{1}{2n + 1}.$$

From this telescoping form, you can see that the nth partial sum is

$$S_n = \left(\frac{1}{1} - \frac{1}{3}\right) + \left(\frac{1}{3} - \frac{1}{5}\right) + \cdots + \left(\frac{1}{2n - 1} - \frac{1}{2n + 1}\right) = 1 - \frac{1}{2n + 1}.$$

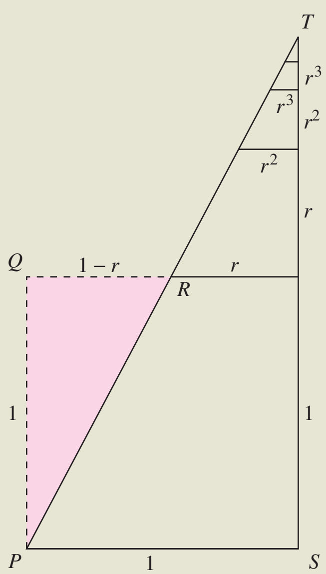
So, the series converges and its sum is 1. That is,

$$\sum_{n=1}^{\infty} \frac{2}{4n^2 - 1} = \lim_{n\to\infty} S_n = \lim_{n\to\infty} \left(1 - \frac{1}{2n + 1}\right) = 1.$$

Try It | Exploration A | Exploration B

EXPLORATION

In "Proof Without Words," by Benjamin G. Klein and Irl C. Bivens, the authors present the following diagram. Explain why the final statement below the diagram is valid. How is this result related to Theorem 9.6?



$\Delta PQR \approx \Delta TSP$

$$1 + r + r^2 + r^3 + \cdots = \frac{1}{1 - r}$$

Exercise taken from "Proof Without Words" by Benjamin G. Klein and Irl C. Bivens, *Mathematics Magazine*, October 1988, by permission of the authors.

Geometric Series

The series given in Example 1(a) is a **geometric series.** In general, the series given by

$$\sum_{n=0}^{\infty} ar^n = a + ar + ar^2 + \cdots + ar^n + \cdots, \quad a \neq 0 \qquad \text{Geometric series}$$

is a **geometric series** with ratio r.

THEOREM 9.6 Convergence of a Geometric Series

A geometric series with ratio r diverges if $|r| \geq 1$. If $0 < |r| < 1$, then the series converges to the sum

$$\sum_{n=0}^{\infty} ar^n = \frac{a}{1 - r}, \quad 0 < |r| < 1.$$

Proof It is easy to see that the series diverges if $r = \pm 1$. If $r \neq \pm 1$, then $S_n = a + ar + ar^2 + \cdots + ar^{n-1}$. Multiplication by r yields

$$rS_n = ar + ar^2 + ar^3 + \cdots + ar^n.$$

Subtracting the second equation from the first produces $S_n - rS_n = a - ar^n$. Therefore, $S_n(1 - r) = a(1 - r^n)$, and the nth partial sum is

$$S_n = \frac{a}{1 - r}(1 - r^n).$$

If $0 < |r| < 1$, it follows that $r^n \to 0$ as $n \to \infty$, and you obtain

$$\lim_{n\to\infty} S_n = \lim_{n\to\infty} \left[\frac{a}{1 - r}(1 - r^n)\right] = \frac{a}{1 - r}\left[\lim_{n\to\infty} (1 - r^n)\right] = \frac{a}{1 - r}$$

which means that the series *converges* and its sum is $a/(1 - r)$. It is left to you to show that the series diverges if $|r| > 1$.

TECHNOLOGY Try using a graphing utility or writing a computer program to compute the sum of the first 20 terms of the sequence in Example 3(a). You should obtain a sum of about 5.999994.

EXAMPLE 3 Convergent and Divergent Geometric Series

a. The geometric series

$$\sum_{n=0}^{\infty} \frac{3}{2^n} = \sum_{n=0}^{\infty} 3\left(\frac{1}{2}\right)^n$$
$$= 3(1) + 3\left(\frac{1}{2}\right) + 3\left(\frac{1}{2}\right)^2 + \cdots$$

has a ratio of $r = \frac{1}{2}$ with $a = 3$. Because $0 < |r| < 1$, the series converges and its sum is

$$S = \frac{a}{1-r} = \frac{3}{1-(1/2)} = 6.$$

b. The geometric series

$$\sum_{n=0}^{\infty} \left(\frac{3}{2}\right)^n = 1 + \frac{3}{2} + \frac{9}{4} + \frac{27}{8} + \cdots$$

has a ratio of $r = \frac{3}{2}$. Because $|r| \geq 1$, the series diverges.

Try It | Exploration A | Exploration B | Exploration C

Exploration D | Exploration E | Technology

The formula for the sum of a geometric series can be used to write a repeating decimal as the ratio of two integers, as demonstrated in the next example.

EXAMPLE 4 A Geometric Series for a Repeating Decimal

Use a geometric series to write $0.\overline{08}$ as the ratio of two integers.

Solution For the repeating decimal $0.\overline{08}$, you can write

$$0.080808\ldots = \frac{8}{10^2} + \frac{8}{10^4} + \frac{8}{10^6} + \frac{8}{10^8} + \cdots$$
$$= \sum_{n=0}^{\infty} \left(\frac{8}{10^2}\right)\left(\frac{1}{10^2}\right)^n.$$

For this series, you have $a = 8/10^2$ and $r = 1/10^2$. So,

$$0.080808\ldots = \frac{a}{1-r} = \frac{8/10^2}{1-(1/10^2)} = \frac{8}{99}.$$

Try dividing 8 by 99 on a calculator to see that it produces $0.\overline{08}$.

Try It | Exploration A | Open Exploration

The convergence of a series is not affected by removal of a finite number of terms from the beginning of the series. For instance, the geometric series

$$\sum_{n=4}^{\infty} \left(\frac{1}{2}\right)^n \quad \text{and} \quad \sum_{n=0}^{\infty} \left(\frac{1}{2}\right)^n$$

both converge. Furthermore, because the sum of the second series is $a/(1-r) = 2$, you can conclude that the sum of the first series is

$$S = 2 - \left[\left(\frac{1}{2}\right)^0 + \left(\frac{1}{2}\right)^1 + \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^3\right]$$
$$= 2 - \frac{15}{8} = \frac{1}{8}.$$

STUDY TIP As you study this chapter, it is important to distinguish between an infinite series and a sequence. A sequence is an ordered collection of numbers

$$a_1, a_2, a_3, \ldots, a_n, \ldots$$

whereas a series is an infinite sum of terms from a sequence

$$a_1 + a_2 + \cdots + a_n + \cdots.$$

The following properties are direct consequences of the corresponding properties of limits of sequences.

THEOREM 9.7 Properties of Infinite Series

If $\Sigma a_n = A$, $\Sigma b_n = B$, and c is a real number, then the following series converge to the indicated sums.

1. $\displaystyle\sum_{n=1}^{\infty} ca_n = cA$
2. $\displaystyle\sum_{n=1}^{\infty} (a_n + b_n) = A + B$
3. $\displaystyle\sum_{n=1}^{\infty} (a_n - b_n) = A - B$

*n*th-Term Test for Divergence

The following theorem states that if a series converges, the limit of its nth term must be 0.

NOTE Be sure you see that the converse of Theorem 9.8 is generally not true. That is, if the sequence $\{a_n\}$ converges to 0, then the series Σa_n may either converge or diverge.

THEOREM 9.8 Limit of *n*th Term of a Convergent Series

If $\displaystyle\sum_{n=1}^{\infty} a_n$ converges, then $\displaystyle\lim_{n\to\infty} a_n = 0$.

Proof Assume that

$$\sum_{n=1}^{\infty} a_n = \lim_{n\to\infty} S_n = L.$$

Then, because $S_n = S_{n-1} + a_n$ and

$$\lim_{n\to\infty} S_n = \lim_{n\to\infty} S_{n-1} = L$$

it follows that

$$\begin{aligned} L = \lim_{n\to\infty} S_n &= \lim_{n\to\infty} (S_{n-1} + a_n) \\ &= \lim_{n\to\infty} S_{n-1} + \lim_{n\to\infty} a_n \\ &= L + \lim_{n\to\infty} a_n \end{aligned}$$

which implies that $\{a_n\}$ converges to 0.

The contrapositive of Theorem 9.8 provides a useful test for *divergence*. This ***n*th-Term Test for Divergence** states that if the limit of the nth term of a series does *not* converge to 0, the series must diverge.

THEOREM 9.9 *n*th-Term Test for Divergence

If $\displaystyle\lim_{n\to\infty} a_n \neq 0$, then $\displaystyle\sum_{n=1}^{\infty} a_n$ diverges.

EXAMPLE 5 Using the *n*th-Term Test for Divergence

a. For the series $\sum_{n=0}^{\infty} 2^n$, you have

$$\lim_{n\to\infty} 2^n = \infty.$$

So, the limit of the nth term is not 0, and the series diverges.

b. For the series $\sum_{n=1}^{\infty} \frac{n!}{2n! + 1}$, you have

$$\lim_{n\to\infty} \frac{n!}{2n! + 1} = \frac{1}{2}.$$

So, the limit of the nth term is not 0, and the series diverges.

c. For the series $\sum_{n=1}^{\infty} \frac{1}{n}$, you have

$$\lim_{n\to\infty} \frac{1}{n} = 0.$$

Because the limit of the nth term is 0, the nth-Term Test for Divergence does *not* apply and you can draw no conclusions about convergence or divergence. (In the next section, you will see that this particular series diverges.)

STUDY TIP The series in Example 5(c) will play an important role in this chapter.

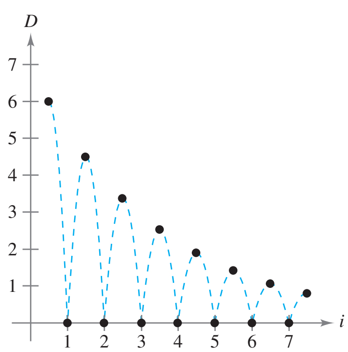
$$\sum_{n=1}^{\infty} \frac{1}{n} = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots$$

You will see that this series diverges even though the nth term approaches 0 as n approaches ∞.

Try It Exploration A

EXAMPLE 6 Bouncing Ball Problem

A ball is dropped from a height of 6 feet and begins bouncing, as shown in Figure 9.7. The height of each bounce is three-fourths the height of the previous bounce. Find the total vertical distance traveled by the ball.



The height of each bounce is three-fourths the height of the preceding bounce.
Figure 9.7

Simulation

Solution When the ball hits the ground for the first time, it has traveled a distance of $D_1 = 6$ feet. For subsequent bounces, let D_i be the distance traveled up and down. For example, D_2 and D_3 are as follows.

$$D_2 = \underbrace{6\left(\tfrac{3}{4}\right)}_{\text{Up}} + \underbrace{6\left(\tfrac{3}{4}\right)}_{\text{Down}} = 12\left(\tfrac{3}{4}\right)$$

$$D_3 = \underbrace{6\left(\tfrac{3}{4}\right)\left(\tfrac{3}{4}\right)}_{\text{Up}} + \underbrace{6\left(\tfrac{3}{4}\right)\left(\tfrac{3}{4}\right)}_{\text{Down}} = 12\left(\tfrac{3}{4}\right)^2$$

By continuing this process, it can be determined that the total vertical distance is

$$\begin{aligned}
D &= 6 + 12\left(\tfrac{3}{4}\right) + 12\left(\tfrac{3}{4}\right)^2 + 12\left(\tfrac{3}{4}\right)^3 + \cdots \\
&= 6 + 12 \sum_{n=0}^{\infty} \left(\tfrac{3}{4}\right)^{n+1} \\
&= 6 + 12\left(\tfrac{3}{4}\right) \sum_{n=0}^{\infty} \left(\tfrac{3}{4}\right)^{n} \\
&= 6 + 9\left(\frac{1}{1 - \frac{3}{4}}\right) \\
&= 6 + 9(4) \\
&= 42 \text{ feet.}
\end{aligned}$$

Try It Exploration A

Section 9.3 The Integral Test and p-Series

- **Use the Integral Test to determine whether an infinite series converges or diverges.**
- **Use properties of p-series and harmonic series.**

The Integral Test

In this and the following section, you will study several convergence tests that apply to series with *positive* terms.

THEOREM 9.10 The Integral Test

If f is positive, continuous, and decreasing for $x \geq 1$ and $a_n = f(n)$, then

$$\sum_{n=1}^{\infty} a_n \quad \text{and} \quad \int_1^{\infty} f(x)\,dx$$

either both converge or both diverge.

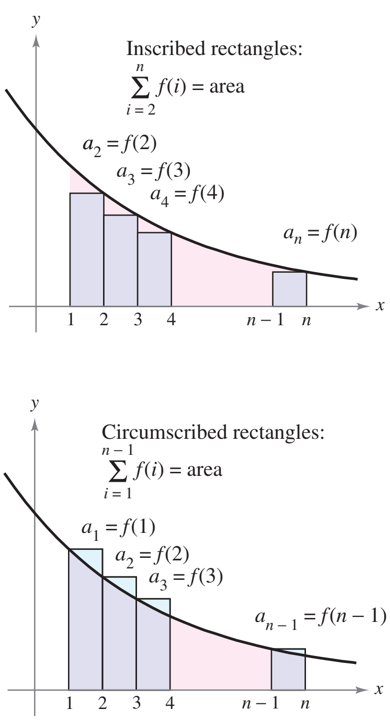


Figure 9.8

Proof Begin by partitioning the interval $[1, n]$ into $n - 1$ unit intervals, as shown in Figure 9.8. The total areas of the inscribed rectangles and the circumscribed rectangles are as follows.

$$\sum_{i=2}^{n} f(i) = f(2) + f(3) + \cdot \cdot \cdot + f(n) \qquad \text{Inscribed area}$$

$$\sum_{i=1}^{n-1} f(i) = f(1) + f(2) + \cdot \cdot \cdot + f(n-1) \qquad \text{Circumscribed area}$$

The exact area under the graph of f from $x = 1$ to $x = n$ lies between the inscribed and circumscribed areas.

$$\sum_{i=2}^{n} f(i) \leq \int_1^n f(x)\,dx \leq \sum_{i=1}^{n-1} f(i)$$

Using the nth partial sum, $S_n = f(1) + f(2) + \cdot \cdot \cdot + f(n)$, you can write this inequality as

$$S_n - f(1) \leq \int_1^n f(x)\,dx \leq S_{n-1}.$$

Now, assuming that $\int_1^{\infty} f(x)\,dx$ converges to L, it follows that for $n \geq 1$

$$S_n - f(1) \leq L \quad \Rightarrow \quad S_n \leq L + f(1).$$

Consequently, $\{S_n\}$ is bounded and monotonic, and by Theorem 9.5 it converges. So, $\Sigma\, a_n$ converges. For the other direction of the proof, assume that the improper integral diverges. Then $\int_1^n f(x)\,dx$ approaches infinity as $n \to \infty$, and the inequality $S_{n-1} \geq \int_1^n f(x)\,dx$ implies that $\{S_n\}$ diverges. So, $\Sigma\, a_n$ diverges.

NOTE Remember that the convergence or divergence of $\Sigma\, a_n$ is not affected by deleting the first N terms. Similarly, if the conditions for the Integral Test are satisfied for all $x \geq N > 1$, you can simply use the integral $\int_N^{\infty} f(x)\,dx$ to test for convergence or divergence. (This is illustrated in Example 4.)

EXAMPLE 1 Using the Integral Test

Apply the Integral Test to the series $\sum_{n=1}^{\infty} \frac{n}{n^2+1}$.

Solution The function $f(x) = x/(x^2+1)$ is positive and continuous for $x \geq 1$. To determine whether f is decreasing, find the derivative.

$$f'(x) = \frac{(x^2+1)(1) - x(2x)}{(x^2+1)^2} = \frac{-x^2+1}{(x^2+1)^2}$$

So, $f'(x) < 0$ for $x > 1$ and it follows that f satisfies the conditions for the Integral Test. You can integrate to obtain

$$\begin{aligned}\int_1^{\infty} \frac{x}{x^2+1}\,dx &= \frac{1}{2}\int_1^{\infty} \frac{2x}{x^2+1}\,dx\\ &= \frac{1}{2}\lim_{b\to\infty}\int_1^{b} \frac{2x}{x^2+1}\,dx\\ &= \frac{1}{2}\lim_{b\to\infty}\Big[\ln(x^2+1)\Big]_1^b\\ &= \frac{1}{2}\lim_{b\to\infty}\left[\ln(b^2+1) - \ln 2\right]\\ &= \infty.\end{aligned}$$

So, the series *diverges*.

Try It | Exploration A | Technology

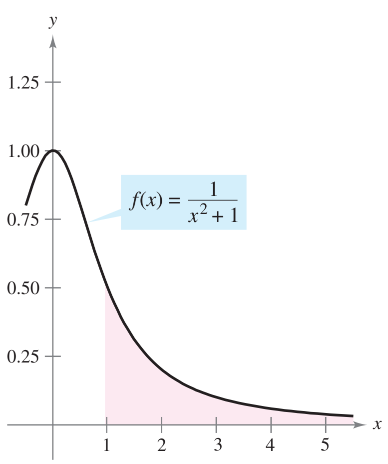
EXAMPLE 2 Using the Integral Test

Apply the Integral Test to the series $\sum_{n=1}^{\infty} \frac{1}{n^2+1}$.

Solution Because $f(x) = 1/(x^2+1)$ satisfies the conditions for the Integral Test (check this), you can integrate to obtain

$$\begin{aligned}\int_1^{\infty} \frac{1}{x^2+1}\,dx &= \lim_{b\to\infty}\int_1^{b} \frac{1}{x^2+1}\,dx\\ &= \lim_{b\to\infty}\Big[\arctan x\Big]_1^b\\ &= \lim_{b\to\infty}(\arctan b - \arctan 1)\\ &= \frac{\pi}{2} - \frac{\pi}{4} = \frac{\pi}{4}.\end{aligned}$$

So, the series *converges* (see Figure 9.9).



Because the improper integral converges, the infinite series also converges.
Figure 9.9

Editable Graph

Try It | Open Exploration

TECHNOLOGY
In Example 2, the fact that the improper integral converges to $\pi/4$ does not imply that the infinite series converges to $\pi/4$. To approximate the sum of the series, you can use the inequality

$$\sum_{n=1}^{N} \frac{1}{n^2+1} \leq \sum_{n=1}^{\infty} \frac{1}{n^2+1} \leq \sum_{n=1}^{N} \frac{1}{n^2+1} + \int_N^{\infty} \frac{1}{x^2+1}\,dx.$$

(See Exercise 60.) The larger the value of N, the better the approximation. For instance, using $N = 200$ produces $1.072 \leq \Sigma 1/(n^2+1) \leq 1.077$.

p-Series and Harmonic Series

In the remainder of this section, you will investigate a second type of series that has a simple arithmetic test for convergence or divergence. A series of the form

$$\sum_{n=1}^{\infty} \frac{1}{n^p} = \frac{1}{1^p} + \frac{1}{2^p} + \frac{1}{3^p} + \cdot \cdot \cdot \qquad p\text{-series}$$

is a ***p*-series,** where p is a positive constant. For $p = 1$, the series

$$\sum_{n=1}^{\infty} \frac{1}{n} = 1 + \frac{1}{2} + \frac{1}{3} + \cdot \cdot \cdot \qquad \text{Harmonic series}$$

is the **harmonic** series. A **general harmonic series** is of the form $\Sigma 1/(an + b)$. In music, strings of the same material, diameter, and tension, whose lengths form a harmonic series, produce harmonic tones.

The Integral Test is convenient for establishing the convergence or divergence of p-series. This is shown in the proof of Theorem 9.11.

HARMONIC SERIES

Pythagoras and his students paid close attention to the development of music as an abstract science. This led to the discovery of the relationship between the tone and the length of the vibrating string. It was observed that the most beautiful musical harmonies corresponded to the simplest ratios of whole numbers. Later mathematicians developed this idea into the harmonic series, where the terms in the harmonic series correspond to the nodes on a vibrating string that produce multiples of the fundamental frequency. For example, $\frac{1}{2}$ is twice the fundamental frequency, $\frac{1}{3}$ is three times the fundamental frequency, and so on.

THEOREM 9.11 Convergence of p-Series

The p-series

$$\sum_{n=1}^{\infty} \frac{1}{n^p} = \frac{1}{1^p} + \frac{1}{2^p} + \frac{1}{3^p} + \frac{1}{4^p} + \cdot \cdot \cdot$$

1. converges if $p > 1$, and
2. diverges if $0 < p \leq 1$.

Proof The proof follows from the Integral Test and from Theorem 8.5, which states that

$$\int_1^{\infty} \frac{1}{x^p}\,dx$$

converges if $p > 1$ and diverges if $0 < p \leq 1$.

EXAMPLE 3 Convergent and Divergent p-Series

Discuss the convergence or divergence of (a) the harmonic series and (b) the p-series with $p = 2$.

Solution

a. From Theorem 9.11, it follows that the harmonic series

$$\sum_{n=1}^{\infty} \frac{1}{n} = \frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \cdot \cdot \cdot \qquad p = 1$$

diverges.

b. From Theorem 9.11, it follows that the p-series

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{1}{1^2} + \frac{1}{2^2} + \frac{1}{3^2} + \cdot \cdot \cdot \qquad p = 2$$

converges.

NOTE The sum of the series in Example 3(b) can be shown to be $\pi^2/6$. (This was proved by Leonhard Euler, but the proof is too difficult to present here.) Be sure you see that the Integral Test does not tell you that the sum of the series is equal to the value of the integral. For instance, the sum of the series in Example 3(b) is

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6} \approx 1.645$$

but the value of the corresponding improper integral is

$$\int_1^{\infty} \frac{1}{x^2}\,dx = 1.$$

Try It Exploration A Exploration B Exploration C

EXAMPLE 4 Testing a Series for Convergence

Determine whether the following series converges or diverges.

$$\sum_{n=2}^{\infty} \frac{1}{n \ln n}$$

Solution This series is similar to the divergent harmonic series. If its terms were larger than those of the harmonic series, you would expect it to diverge. However, because its terms are smaller, you are not sure what to expect. The function $f(x) = 1/(x \ln x)$ is positive and continuous for $x \geq 2$. To determine whether f is decreasing, first rewrite f as $f(x) = (x \ln x)^{-1}$ and then find its derivative.

$$f'(x) = (-1)(x \ln x)^{-2}(1 + \ln x) = -\frac{1 + \ln x}{x^2(\ln x)^2}$$

So, $f'(x) < 0$ for $x > 2$ and it follows that f satisfies the conditions for the Integral Test.

$$\begin{aligned}
\int_2^{\infty} \frac{1}{x \ln x}\, dx &= \int_2^{\infty} \frac{1/x}{\ln x}\, dx \\
&= \lim_{b\to\infty} \Big[\ln(\ln x) \Big]_2^b \\
&= \lim_{b\to\infty} \big[\ln(\ln b) - \ln(\ln 2)\big] = \infty
\end{aligned}$$

The series diverges.

Try It

Exploration A

NOTE The infinite series in Example 4 diverges very slowly. For instance, the sum of the first 10 terms is approximately 1.6878196, whereas the sum of the first 100 terms is just slightly larger: 2.3250871. In fact, the sum of the first 10,000 terms is approximately 3.015021704. You can see that although the infinite series "adds up to infinity," it does so very slowly.

Section 9.4

Comparisons of Series

- **Use the Direct Comparison Test to determine whether a series converges or diverges.**
- **Use the Limit Comparison Test to determine whether a series converges or diverges.**

Direct Comparison Test

For the convergence tests developed so far, the terms of the series have to be fairly simple and the series must have special characteristics in order for the convergence tests to be applied. A slight deviation from these special characteristics can make a test nonapplicable. For example, in the following pairs, the second series cannot be tested by the same convergence test as the first series even though it is similar to the first.

1. $\displaystyle\sum_{n=0}^{\infty} \frac{1}{2^n}$ is geometric, but $\displaystyle\sum_{n=0}^{\infty} \frac{n}{2^n}$ is not.

2. $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n^3}$ is a p-series, but $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n^3 + 1}$ is not.

3. $a_n = \dfrac{n}{(n^2 + 3)^2}$ is easily integrated, but $b_n = \dfrac{n^2}{(n^2 + 3)^2}$ is not.

In this section you will study two additional tests for positive-term series. These two tests greatly expand the variety of series you are able to test for convergence or divergence. They allow you to *compare* a series having complicated terms with a simpler series whose convergence or divergence is known.

THEOREM 9.12 Direct Comparison Test

Let $0 < a_n \leq b_n$ for all n.

1. If $\displaystyle\sum_{n=1}^{\infty} b_n$ converges, then $\displaystyle\sum_{n=1}^{\infty} a_n$ converges.

2. If $\displaystyle\sum_{n=1}^{\infty} a_n$ diverges, then $\displaystyle\sum_{n=1}^{\infty} b_n$ diverges.

Proof To prove the first property, let $L = \displaystyle\sum_{n=1}^{\infty} b_n$ and let

$$S_n = a_1 + a_2 + \cdot \cdot \cdot + a_n.$$

Because $0 < a_n \leq b_n$, the sequence $S_1, S_2, S_3, \ldots$ is nondecreasing and bounded above by L; so, it must converge. Because

$$\lim_{n\to\infty} S_n = \sum_{n=1}^{\infty} a_n$$

it follows that $\Sigma\, a_n$ converges. The second property is logically equivalent to the first.

NOTE As stated, the Direct Comparison Test requires that $0 < a_n \leq b_n$ for all n. Because the convergence of a series is not dependent on its first several terms, you could modify the test to require only that $0 < a_n \leq b_n$ for all n greater than some integer N.

EXAMPLE 1 Using the Direct Comparison Test

Determine the convergence or divergence of

$$\sum_{n=1}^{\infty} \frac{1}{2+3^n}.$$

Solution This series resembles

$$\sum_{n=1}^{\infty} \frac{1}{3^n}. \qquad \text{Convergent geometric series}$$

Term-by-term comparison yields

$$a_n = \frac{1}{2+3^n} < \frac{1}{3^n} = b_n, \quad n \geq 1.$$

So, by the Direct Comparison Test, the series converges.

Try It Exploration A Technology

EXAMPLE 2 Using the Direct Comparison Test

Determine the convergence or divergence of

$$\sum_{n=1}^{\infty} \frac{1}{2+\sqrt{n}}.$$

Solution This series resembles

$$\sum_{n=1}^{\infty} \frac{1}{n^{1/2}}. \qquad \text{Divergent } p\text{-series}$$

Term-by-term comparison yields

$$\frac{1}{2+\sqrt{n}} \leq \frac{1}{\sqrt{n}}, \quad n \geq 1$$

which *does not* meet the requirements for divergence. (Remember that if term-by-term comparison reveals a series that is *smaller* than a divergent series, the Direct Comparison Test tells you nothing.) Still expecting the series to diverge, you can compare the given series with

$$\sum_{n=1}^{\infty} \frac{1}{n}. \qquad \text{Divergent harmonic series}$$

In this case, term-by-term comparison yields

$$a_n = \frac{1}{n} \leq \frac{1}{2+\sqrt{n}} = b_n, \quad n \geq 4$$

and, by the Direct Comparison Test, the given series diverges.

NOTE To verify the last inequality in Example 2, try showing that $2 + \sqrt{n} \leq n$ whenever $n \geq 4$.

Try It Exploration A Open Exploration

Remember that both parts of the Direct Comparison Test require that $0 < a_n \leq b_n$. Informally, the test says the following about the two series with nonnegative terms.

1. If the "larger" series converges, the "smaller" series must also converge.
2. If the "smaller" series diverges, the "larger" series must also diverge.

Limit Comparison Test

Often a given series closely resembles a p-series or a geometric series, yet you cannot establish the term-by-term comparison necessary to apply the Direct Comparison Test. Under these circumstances you may be able to apply a second comparison test, called the **Limit Comparison Test.**

THEOREM 9.13 Limit Comparison Test

Suppose that $a_n > 0$, $b_n > 0$, and

$$\lim_{n\to\infty}\left(\frac{a_n}{b_n}\right) = L$$

where L is *finite and positive*. Then the two series $\Sigma\, a_n$ and $\Sigma\, b_n$ either both converge or both diverge.

NOTE As with the Direct Comparison Test, the Limit Comparison Test could be modified to require only that a_n and b_n be positive for all n greater than some integer N.

Proof Because $a_n > 0$, $b_n > 0$, and

$$\lim_{n\to\infty}\left(\frac{a_n}{b_n}\right) = L$$

there exists $N > 0$ such that

$$0 < \frac{a_n}{b_n} < L + 1, \quad \text{for } n \geq N.$$

This implies that

$$0 < a_n < (L + 1)b_n.$$

So, by the Direct Comparison Test, the convergence of $\Sigma\, b_n$ implies the convergence of $\Sigma\, a_n$. Similarly, the fact that

$$\lim_{n\to\infty}\left(\frac{b_n}{a_n}\right) = \frac{1}{L}$$

can be used to show that the convergence of $\Sigma\, a_n$ implies the convergence of $\Sigma\, b_n$.

EXAMPLE 3 Using the Limit Comparison Test

Show that the following general harmonic series diverges.

$$\sum_{n=1}^{\infty} \frac{1}{an + b}, \quad a > 0, \quad b > 0$$

Solution By comparison with

$$\sum_{n=1}^{\infty} \frac{1}{n} \qquad \text{Divergent harmonic series}$$

you have

$$\lim_{n\to\infty} \frac{1/(an + b)}{1/n} = \lim_{n\to\infty} \frac{n}{an + b} = \frac{1}{a}.$$

Because this limit is greater than 0, you can conclude from the Limit Comparison Test that the given series diverges.

Try It Exploration A Technology

The Limit Comparison Test works well for comparing a "messy" algebraic series with a p-series. In choosing an appropriate p-series, you must choose one with an nth term of the same magnitude as the nth term of the given series.

Given Series	*Comparison Series*	*Conclusion*
$\displaystyle\sum_{n=1}^{\infty} \frac{1}{3n^2 - 4n + 5}$	$\displaystyle\sum_{n=1}^{\infty} \frac{1}{n^2}$	Both series converge.
$\displaystyle\sum_{n=1}^{\infty} \frac{1}{\sqrt{3n - 2}}$	$\displaystyle\sum_{n=1}^{\infty} \frac{1}{\sqrt{n}}$	Both series diverge.
$\displaystyle\sum_{n=1}^{\infty} \frac{n^2 - 10}{4n^5 + n^3}$	$\displaystyle\sum_{n=1}^{\infty} \frac{n^2}{n^5} = \sum_{n=1}^{\infty} \frac{1}{n^3}$	Both series converge.

In other words, when choosing a series for comparison, you can disregard all but the *highest powers of n* in both the numerator and the denominator.

EXAMPLE 4 Using the Limit Comparison Test

Determine the convergence or divergence of

$$\sum_{n=1}^{\infty} \frac{\sqrt{n}}{n^2 + 1}.$$

Solution Disregarding all but the highest powers of n in the numerator and the denominator, you can compare the series with

$$\sum_{n=1}^{\infty} \frac{\sqrt{n}}{n^2} = \sum_{n=1}^{\infty} \frac{1}{n^{3/2}}.$$ Convergent p-series

Because

$$\begin{aligned} \lim_{n\to\infty} \frac{a_n}{b_n} &= \lim_{n\to\infty} \left(\frac{\sqrt{n}}{n^2 + 1}\right)\left(\frac{n^{3/2}}{1}\right) \\ &= \lim_{n\to\infty} \frac{n^2}{n^2 + 1} = 1 \end{aligned}$$

you can conclude by the Limit Comparison Test that the given series converges.

Try It **Exploration A**

EXAMPLE 5 Using the Limit Comparison Test

Determine the convergence or divergence of

$$\sum_{n=1}^{\infty} \frac{n2^n}{4n^3 + 1}.$$

Solution A reasonable comparison would be with the series

$$\sum_{n=1}^{\infty} \frac{2^n}{n^2}.$$ Divergent series

Note that this series diverges by the nth-Term Test. From the limit

$$\begin{aligned} \lim_{n\to\infty} \frac{a_n}{b_n} &= \lim_{n\to\infty} \left(\frac{n2^n}{4n^3 + 1}\right)\left(\frac{n^2}{2^n}\right) \\ &= \lim_{n\to\infty} \frac{1}{4 + (1/n^3)} = \frac{1}{4} \end{aligned}$$

you can conclude that the given series diverges.

Try It **Exploration A**

Section 9.5 Alternating Series

- **Use the Alternating Series Test to determine whether an infinite series converges.**
- **Use the Alternating Series Remainder to approximate the sum of an alternating series.**
- **Classify a convergent series as absolutely or conditionally convergent.**
- **Rearrange an infinite series to obtain a different sum.**

Alternating Series

So far, most series you have dealt with have had positive terms. In this section and the following section, you will study series that contain both positive and negative terms. The simplest such series is an **alternating series,** whose terms alternate in sign. For example, the geometric series

$$\sum_{n=0}^{\infty}\left(-\frac{1}{2}\right)^n = \sum_{n=0}^{\infty}(-1)^n\frac{1}{2^n}$$
$$= 1 - \frac{1}{2} + \frac{1}{4} - \frac{1}{8} + \frac{1}{16} - \cdots$$

is an *alternating geometric series* with $r = -\frac{1}{2}$. Alternating series occur in two ways: either the odd terms are negative or the even terms are negative.

> **THEOREM 9.14 Alternating Series Test**
>
> Let $a_n > 0$. The alternating series
>
> $$\sum_{n=1}^{\infty}(-1)^n a_n \quad \text{and} \quad \sum_{n=1}^{\infty}(-1)^{n+1} a_n$$
>
> converge if the following two conditions are met.
>
> **1.** $\lim_{n\to\infty} a_n = 0$ **2.** $a_{n+1} \le a_n$, for all n

Proof Consider the alternating series $\Sigma\,(-1)^{n+1} a_n$. For this series, the partial sum (where $2n$ is even)

$$S_{2n} = (a_1 - a_2) + (a_3 - a_4) + (a_5 - a_6) + \cdots + (a_{2n-1} - a_{2n})$$

has all nonnegative terms, and therefore $\{S_{2n}\}$ is a nondecreasing sequence. But you can also write

$$S_{2n} = a_1 - (a_2 - a_3) - (a_4 - a_5) - \cdots - (a_{2n-2} - a_{2n-1}) - a_{2n}$$

which implies that $S_{2n} \le a_1$ for every integer n. So, $\{S_{2n}\}$ is a bounded, nondecreasing sequence that converges to some value L. Because $S_{2n-1} - a_{2n} = S_{2n}$ and $a_{2n} \to 0$, you have

$$\lim_{n\to\infty} S_{2n-1} = \lim_{n\to\infty} S_{2n} + \lim_{n\to\infty} a_{2n}$$
$$= L + \lim_{n\to\infty} a_{2n} = L.$$

Because both S_{2n} and S_{2n-1} converge to the same limit L, it follows that $\{S_n\}$ also converges to L. Consequently, the given alternating series converges.

NOTE The second condition in the Alternating Series Test can be modified to require only that $0 < a_{n+1} \le a_n$ for all n greater than some integer N.

EXAMPLE 1 Using the Alternating Series Test

NOTE The series in Example 1 is called the *alternating harmonic series*—more is said about this series in Example 7.

Determine the convergence or divergence of $\displaystyle\sum_{n=1}^{\infty} (-1)^{n+1}\frac{1}{n}$.

Solution Note that $\displaystyle\lim_{n\to\infty} a_n = \lim_{n\to\infty}\frac{1}{n} = 0$. So, the first condition of Theorem 9.14 is satisfied. Also note that the second condition of Theorem 9.14 is satisfied because

$$a_{n+1} = \frac{1}{n+1} \le \frac{1}{n} = a_n$$

for all n. So, applying the Alternating Series Test, you can conclude that the series converges.

Try It | Exploration A | Technology

EXAMPLE 2 Using the Alternating Series Test

Determine the convergence or divergence of $\displaystyle\sum_{n=1}^{\infty} \frac{n}{(-2)^{n-1}}$.

Solution To apply the Alternating Series Test, note that, for $n \ge 1$,

$$\begin{aligned}
\frac{1}{2} &\le \frac{n}{n+1}\\
\frac{2^{n-1}}{2^n} &\le \frac{n}{n+1}\\
(n+1)2^{n-1} &\le n2^n\\
\frac{n+1}{2^n} &\le \frac{n}{2^{n-1}}.
\end{aligned}$$

So, $a_{n+1} = (n+1)/2^n \le n/2^{n-1} = a_n$ for all n. Furthermore, by L'Hôpital's Rule,

$$\lim_{x\to\infty}\frac{x}{2^{x-1}} = \lim_{x\to\infty}\frac{1}{2^{x-1}(\ln 2)} = 0 \quad\Rightarrow\quad \lim_{n\to\infty}\frac{n}{2^{n-1}} = 0.$$

Therefore, by the Alternating Series Test, the series converges.

Try It | Exploration A | Exploration B

EXAMPLE 3 Cases for Which the Alternating Series Test Fails

NOTE In Example 3(a), remember that whenever a series does not pass the first condition of the Alternating Series Test, you can use the nth-Term Test for Divergence to conclude that the series diverges.

a. The alternating series

$$\sum_{n=1}^{\infty}\frac{(-1)^{n+1}(n+1)}{n} = \frac{2}{1} - \frac{3}{2} + \frac{4}{3} - \frac{5}{4} + \frac{6}{5} - \cdots$$

passes the second condition of the Alternating Series Test because $a_{n+1} \le a_n$ for all n. You cannot apply the Alternating Series Test, however, because the series does not pass the first condition. In fact, the series diverges.

b. The alternating series

$$\frac{2}{1} - \frac{1}{1} + \frac{2}{2} - \frac{1}{2} + \frac{2}{3} - \frac{1}{3} + \frac{2}{4} - \frac{1}{4} + \cdots$$

passes the first condition because a_n approaches 0 as $n \to \infty$. You cannot apply the Alternating Series Test, however, because the series does not pass the second condition. To conclude that the series diverges, you can argue that S_{2N} equals the Nth partial sum of the divergent harmonic series. This implies that the sequence of partial sums diverges. So, the series diverges.

Try It | Exploration A

Alternating Series Remainder

For a convergent alternating series, the partial sum S_N can be a useful approximation for the sum S of the series. The error involved in using $S \approx S_N$ is the remainder $R_N = S - S_N$.

THEOREM 9.15 Alternating Series Remainder

If a convergent alternating series satisfies the condition $a_{n+1} \le a_n$, then the absolute value of the remainder R_N involved in approximating the sum S by S_N is less than (or equal to) the first neglected term. That is,

$$|S - S_N| = |R_N| \le a_{N+1}.$$

Proof The series obtained by deleting the first N terms of the given series satisfies the conditions of the Alternating Series Test and has a sum of R_N.

$$\begin{aligned}
R_N = S - S_N &= \sum_{n=1}^{\infty} (-1)^{n+1} a_n - \sum_{n=1}^{N} (-1)^{n+1} a_n \\
&= (-1)^N a_{N+1} + (-1)^{N+1} a_{N+2} + (-1)^{N+2} a_{N+3} + \cdots \\
&= (-1)^N (a_{N+1} - a_{N+2} + a_{N+3} - \cdots) \\
|R_N| &= a_{N+1} - a_{N+2} + a_{N+3} - a_{N+4} + a_{N+5} - \cdots \\
&= a_{N+1} - (a_{N+2} - a_{N+3}) - (a_{N+4} - a_{N+5}) - \cdots \le a_{N+1}
\end{aligned}$$

Consequently, $|S - S_N| = |R_N| \le a_{N+1}$, which establishes the theorem.

EXAMPLE 4 Approximating the Sum of an Alternating Series

Approximate the sum of the following series by its first six terms.

$$\sum_{n=1}^{\infty} (-1)^{n+1} \left(\frac{1}{n!}\right) = \frac{1}{1!} - \frac{1}{2!} + \frac{1}{3!} - \frac{1}{4!} + \frac{1}{5!} - \frac{1}{6!} + \cdots$$

Solution The series converges by the Alternating Series Test because

$$\frac{1}{(n+1)!} \le \frac{1}{n!} \quad \text{and} \quad \lim_{n\to\infty} \frac{1}{n!} = 0.$$

The sum of the first six terms is

$$S_6 = 1 - \frac{1}{2} + \frac{1}{6} - \frac{1}{24} + \frac{1}{120} - \frac{1}{720} = \frac{91}{144} \approx 0.63194$$

and, by the Alternating Series Remainder, you have

$$|S - S_6| = |R_6| \le a_7 = \frac{1}{5040} \approx 0.0002.$$

So, the sum S lies between $0.63194 - 0.0002$ and $0.63194 + 0.0002$, and you have

$$0.63174 \le S \le 0.63214.$$

TECHNOLOGY Later, in Section 9.10, you will be able to show that the series in Example 4 converges to

$$\frac{e-1}{e} \approx 0.63212.$$

For now, try using a computer to obtain an approximation of the sum of the series. How many terms do you need to obtain an approximation that is within 0.00001 unit of the actual sum?

Try It | Exploration A | Open Exploration

Absolute and Conditional Convergence

Occasionally, a series may have both positive and negative terms and not be an alternating series. For instance, the series

$$\sum_{n=1}^{\infty} \frac{\sin n}{n^2} = \frac{\sin 1}{1} + \frac{\sin 2}{4} + \frac{\sin 3}{9} + \cdots$$

has both positive and negative terms, yet it is not an alternating series. One way to obtain some information about the convergence of this series is to investigate the convergence of the series

$$\sum_{n=1}^{\infty} \left|\frac{\sin n}{n^2}\right|.$$

By direct comparison, you have $|\sin n| \le 1$ for all n, so

$$\left|\frac{\sin n}{n^2}\right| \le \frac{1}{n^2}, \quad n \ge 1.$$

Therefore, by the Direct Comparison Test, the series $\sum \left|\frac{\sin n}{n^2}\right|$ converges. The next theorem tells you that the original series also converges.

THEOREM 9.16 Absolute Convergence

If the series $\Sigma|a_n|$ converges, then the series $\Sigma\, a_n$ also converges.

Proof Because $0 \le a_n + |a_n| \le 2|a_n|$ for all n, the series

$$\sum_{n=1}^{\infty} (a_n + |a_n|)$$

converges by comparison with the convergent series

$$\sum_{n=1}^{\infty} 2|a_n|.$$

Furthermore, because $a_n = (a_n + |a_n|) - |a_n|$, you can write

$$\sum_{n=1}^{\infty} a_n = \sum_{n=1}^{\infty} (a_n + |a_n|) - \sum_{n=1}^{\infty} |a_n|$$

where both series on the right converge. So, it follows that $\Sigma\, a_n$ converges.

The converse of Theorem 9.16 is not true. For instance, the **alternating harmonic series**

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} = \frac{1}{1} - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots$$

converges by the Alternating Series Test. Yet the harmonic series diverges. This type of convergence is called **conditional.**

Definitions of Absolute and Conditional Convergence

1. $\Sigma\, a_n$ is **absolutely convergent** if $\Sigma\, |a_n|$ converges.
2. $\Sigma\, a_n$ is **conditionally convergent** if $\Sigma\, a_n$ converges but $\Sigma\, |a_n|$ diverges.

EXAMPLE 5 Absolute and Conditional Convergence

Determine whether each of the series is convergent or divergent. Classify any convergent series as absolutely or conditionally convergent.

a. $\displaystyle\sum_{n=0}^{\infty} \frac{(-1)^n n!}{2^n} = \frac{0!}{2^0} - \frac{1!}{2^1} + \frac{2!}{2^2} - \frac{3!}{2^3} + \cdots$

b. $\displaystyle\sum_{n=1}^{\infty} \frac{(-1)^n}{\sqrt{n}} = -\frac{1}{\sqrt{1}} + \frac{1}{\sqrt{2}} - \frac{1}{\sqrt{3}} + \frac{1}{\sqrt{4}} - \cdots$

Solution

a. By the nth-Term Test for Divergence, you can conclude that this series diverges.

b. The given series can be shown to be convergent by the Alternating Series Test. Moreover, because the p-series

$$\sum_{n=1}^{\infty} \left|\frac{(-1)^n}{\sqrt{n}}\right| = \frac{1}{\sqrt{1}} + \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{3}} + \frac{1}{\sqrt{4}} + \cdots$$

diverges, the given series is *conditionally* convergent.

Try It **Exploration A** **Exploration B**

EXAMPLE 6 Absolute and Conditional Convergence

Determine whether each of the series is convergent or divergent. Classify any convergent series as absolutely or conditionally convergent.

a. $\displaystyle\sum_{n=1}^{\infty} \frac{(-1)^{n(n+1)/2}}{3^n} = -\frac{1}{3} - \frac{1}{9} + \frac{1}{27} + \frac{1}{81} - \cdots$

b. $\displaystyle\sum_{n=1}^{\infty} \frac{(-1)^n}{\ln(n+1)} = -\frac{1}{\ln 2} + \frac{1}{\ln 3} - \frac{1}{\ln 4} + \frac{1}{\ln 5} - \cdots$

Solution

a. This is *not* an alternating series. However, because

$$\sum_{n=1}^{\infty} \left|\frac{(-1)^{n(n+1)/2}}{3^n}\right| = \sum_{n=1}^{\infty} \frac{1}{3^n}$$

is a convergent geometric series, you can apply Theorem 9.16 to conclude that the given series is *absolutely* convergent (and therefore convergent).

b. In this case, the Alternating Series Test indicates that the given series converges. However, the series

$$\sum_{n=1}^{\infty} \left|\frac{(-1)^n}{\ln(n+1)}\right| = \frac{1}{\ln 2} + \frac{1}{\ln 3} + \frac{1}{\ln 4} + \cdots$$

diverges by direct comparison with the terms of the harmonic series. Therefore, the given series is *conditionally* convergent.

Try It **Exploration A**

Rearrangement of Series

A finite sum such as $(1 + 3 - 2 + 5 - 4)$ can be rearranged without changing the value of the sum. This is not necessarily true of an infinite series—it depends on whether the series is absolutely convergent (every rearrangement has the same sum) or conditionally convergent.

EXAMPLE 7 Rearrangement of a Series

FOR FURTHER INFORMATION Georg Friedrich Riemann (1826–1866) proved that if $\Sigma\, a_n$ is conditionally convergent and S is any real number, the terms of the series can be rearranged to converge to S. For more on this topic, see the article "Riemann's Rearrangement Theorem" by Stewart Galanor in *Mathematics Teacher*.

MathArticle

The alternating harmonic series converges to ln 2. That is,

$$\sum_{n=1}^{\infty} (-1)^{n+1}\frac{1}{n} = \frac{1}{1} - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots = \ln 2.$$

(See Exercise 49, Section 9.10.)

Rearrange the series to produce a different sum.

Solution Consider the following rearrangement.

$$\begin{aligned}
1 &- \frac{1}{2} - \frac{1}{4} + \frac{1}{3} - \frac{1}{6} - \frac{1}{8} + \frac{1}{5} - \frac{1}{10} - \frac{1}{12} + \frac{1}{7} - \frac{1}{14} - \cdots \\
&= \left(1 - \frac{1}{2}\right) - \frac{1}{4} + \left(\frac{1}{3} - \frac{1}{6}\right) - \frac{1}{8} + \left(\frac{1}{5} - \frac{1}{10}\right) - \frac{1}{12} + \left(\frac{1}{7} - \frac{1}{14}\right) - \cdots \\
&= \frac{1}{2} - \frac{1}{4} + \frac{1}{6} - \frac{1}{8} + \frac{1}{10} - \frac{1}{12} + \frac{1}{14} - \cdots \\
&= \frac{1}{2}\left(1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \frac{1}{7} - \cdots\right) = \frac{1}{2}(\ln 2)
\end{aligned}$$

By rearranging the terms, you obtain a sum that is half the original sum.

Try It Exploration A

Section 9.6 The Ratio and Root Tests

- **Use the Ratio Test to determine whether a series converges or diverges.**
- **Use the Root Test to determine whether a series converges or diverges.**
- **Review the tests for convergence and divergence of an infinite series.**

EXPLORATION

Writing a Series One of the following conditions guarantees that a series will diverge, two conditions guarantee that a series will converge, and one has no guarantee—the series can either converge or diverge. Which is which? Explain your reasoning.

a. $\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| = 0$

b. $\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| = \frac{1}{2}$

c. $\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| = 1$

d. $\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| = 2$

The Ratio Test

This section begins with a test for absolute convergence—the **Ratio Test.**

THEOREM 9.17 Ratio Test

Let $\Sigma\, a_n$ be a series with nonzero terms.

1. $\Sigma\, a_n$ converges absolutely if $\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| < 1$.

2. $\Sigma\, a_n$ diverges if $\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| > 1$ or $\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| = \infty$.

3. The Ratio Test is inconclusive if $\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| = 1$.

Proof To prove Property 1, assume that

$$\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| = r < 1$$

and choose R such that $0 \le r < R < 1$. By the definition of the limit of a sequence, there exists some $N > 0$ such that $|a_{n+1}/a_n| < R$ for all $n > N$. Therefore, you can write the following inequalities.

$$\begin{aligned}
|a_{N+1}| &< |a_N|R \\
|a_{N+2}| &< |a_{N+1}|R < |a_N|R^2 \\
|a_{N+3}| &< |a_{N+2}|R < |a_{N+1}|R^2 < |a_N|R^3 \\
&\vdots
\end{aligned}$$

The geometric series $\Sigma\, |a_N|R^n = |a_N|R + |a_N|R^2 + \cdots + |a_N|R^n + \cdots$ converges, and so, by the Direct Comparison Test, the series

$$\sum_{n=1}^{\infty} |a_{N+n}| = |a_{N+1}| + |a_{N+2}| + \cdots + |a_{N+n}| + \cdots$$

also converges. This in turn implies that the series $\Sigma\, |a_n|$ converges, because discarding a finite number of terms $(n = N - 1)$ does not affect convergence. Consequently, by Theorem 9.16, the series $\Sigma\, a_n$ converges absolutely. The proof of Property 2 is similar and is left as an exercise (see Exercise 98).

NOTE The fact that the Ratio Test is inconclusive when $|a_{n+1}/a_n| \to 1$ can be seen by comparing the two series $\Sigma\,(1/n)$ and $\Sigma\,(1/n^2)$. The first series diverges and the second one converges, but in both cases

$$\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| = 1.$$

Although the Ratio Test is not a cure for all ills related to tests for convergence, it is particularly useful for series that *converge rapidly*. Series involving factorials or exponentials are frequently of this type.

EXAMPLE 1 Using the Ratio Test

Determine the convergence or divergence of

$$\sum_{n=0}^{\infty} \frac{2^n}{n!}.$$

Solution Because $a_n = 2^n/n!$, you can write the following.

$$\begin{aligned}
\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| &= \lim_{n\to\infty}\left[\frac{2^{n+1}}{(n+1)!} \div \frac{2^n}{n!}\right] \\
&= \lim_{n\to\infty}\left[\frac{2^{n+1}}{(n+1)!} \cdot \frac{n!}{2^n}\right] \\
&= \lim_{n\to\infty} \frac{2}{n+1} \\
&= 0
\end{aligned}$$

Therefore, the series converges.

STUDY TIP A step frequently used in applications of the Ratio Test involves simplifying quotients of factorials. In Example 1, for instance, notice that

$$\frac{n!}{(n+1)!} = \frac{n!}{(n+1)n!} = \frac{1}{n+1}.$$

Try It Exploration A

EXAMPLE 2 Using the Ratio Test

Determine whether each series converges or diverges.

a. $\displaystyle\sum_{n=0}^{\infty} \frac{n^2 2^{n+1}}{3^n}$ **b.** $\displaystyle\sum_{n=1}^{\infty} \frac{n^n}{n!}$

Solution

a. This series converges because the limit of $|a_{n+1}/a_n|$ is less than 1.

$$\begin{aligned}
\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| &= \lim_{n\to\infty}\left[(n+1)^2\left(\frac{2^{n+2}}{3^{n+1}}\right)\left(\frac{3^n}{n^2 2^{n+1}}\right)\right] \\
&= \lim_{n\to\infty} \frac{2(n+1)^2}{3n^2} \\
&= \frac{2}{3} < 1
\end{aligned}$$

b. This series diverges because the limit of $|a_{n+1}/a_n|$ is greater than 1.

$$\begin{aligned}
\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| &= \lim_{n\to\infty}\left[\frac{(n+1)^{n+1}}{(n+1)!}\left(\frac{n!}{n^n}\right)\right] \\
&= \lim_{n\to\infty}\left[\frac{(n+1)^{n+1}}{(n+1)}\left(\frac{1}{n^n}\right)\right] \\
&= \lim_{n\to\infty} \frac{(n+1)^n}{n^n} \\
&= \lim_{n\to\infty}\left(1 + \frac{1}{n}\right)^n \\
&= e > 1
\end{aligned}$$

Try It Exploration A

EXAMPLE 3 A Failure of the Ratio Test

Determine the convergence or divergence of $\displaystyle\sum_{n=1}^{\infty} (-1)^n \frac{\sqrt{n}}{n+1}$.

Solution The limit of $|a_{n+1}/a_n|$ is equal to 1.

$$\begin{aligned}\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| &= \lim_{n\to\infty}\left[\left(\frac{\sqrt{n+1}}{n+2}\right)\left(\frac{n+1}{\sqrt{n}}\right)\right]\\ &= \lim_{n\to\infty}\left[\sqrt{\frac{n+1}{n}}\left(\frac{n+1}{n+2}\right)\right]\\ &= \sqrt{1}(1)\\ &= 1\end{aligned}$$

NOTE The Ratio Test is also inconclusive for any p-series.

So, the Ratio Test is inconclusive. To determine whether the series converges, you need to try a different test. In this case, you can apply the Alternating Series Test. To show that $a_{n+1} \le a_n$, let

$$f(x) = \frac{\sqrt{x}}{x+1}.$$

Then the derivative is

$$f'(x) = \frac{-x+1}{2\sqrt{x}(x+1)^2}.$$

Because the derivative is negative for $x > 1$, you know that f is a decreasing function. Also, by L'Hôpital's Rule,

$$\begin{aligned}\lim_{x\to\infty}\frac{\sqrt{x}}{x+1} &= \lim_{x\to\infty}\frac{1/\left(2\sqrt{x}\right)}{1}\\ &= \lim_{x\to\infty}\frac{1}{2\sqrt{x}}\\ &= 0.\end{aligned}$$

Therefore, by the Alternating Series Test, the series converges.

Try It **Exploration A** **Open Exploration**

The series in Example 3 is *conditionally convergent*. This follows from the fact that the series

$$\sum_{n=1}^{\infty} |a_n|$$

diverges $\left(\text{by the Limit Comparison Test with } \Sigma\, 1/\sqrt{n}\right)$, but the series

$$\sum_{n=1}^{\infty} a_n$$

converges.

TECHNOLOGY A computer or programmable calculator can reinforce the conclusion that the series in Example 3 converges *conditionally*. By adding the first 100 terms of the series, you obtain a sum of about -0.2. (The sum of the first 100 terms of the series $\Sigma\,|a_n|$ is about 17.)

The Root Test

The next test for convergence or divergence of series works especially well for series involving nth powers. The proof of this theorem is similar to that given for the Ratio Test, and is left as an exercise (see Exercise 99).

THEOREM 9.18 Root Test

Let $\Sigma\, a_n$ be a series.

1. $\Sigma\, a_n$ converges absolutely if $\lim_{n\to\infty} \sqrt[n]{|a_n|} < 1$.
2. $\Sigma\, a_n$ diverges if $\lim_{n\to\infty} \sqrt[n]{|a_n|} > 1$ or $\lim_{n\to\infty} \sqrt[n]{|a_n|} = \infty$.
3. The Root Test is inconclusive if $\lim_{n\to\infty} \sqrt[n]{|a_n|} = 1$.

EXAMPLE 4 Using the Root Test

Determine the convergence or divergence of

$$\sum_{n=1}^{\infty} \frac{e^{2n}}{n^n}.$$

Solution You can apply the Root Test as follows.

$$\begin{aligned}
\lim_{n\to\infty} \sqrt[n]{|a_n|} &= \lim_{n\to\infty} \sqrt[n]{\frac{e^{2n}}{n^n}} \\
&= \lim_{n\to\infty} \frac{e^{2n/n}}{n^{n/n}} \\
&= \lim_{n\to\infty} \frac{e^2}{n} \\
&= 0 < 1
\end{aligned}$$

NOTE The Root Test is always inconclusive for any p-series.

Because this limit is less than 1, you can conclude that the series converges absolutely (and therefore converges).

Try It

Exploration A

FOR FURTHER INFORMATION For more information on the usefulness of the Root Test, see the article "N! and the Root Test" by Charles C. Mumma II in *The American Mathematical Monthly*.

MathArticle

To see the usefulness of the Root Test for the series in Example 4, try applying the Ratio Test to that series. When you do this, you obtain the following.

$$\begin{aligned}
\lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| &= \lim_{n\to\infty} \left[\frac{e^{2(n+1)}}{(n+1)^{n+1}} \div \frac{e^{2n}}{n^n}\right] \\
&= \lim_{n\to\infty} e^2 \frac{n^n}{(n+1)^{n+1}} \\
&= \lim_{n\to\infty} e^2 \left(\frac{n}{n+1}\right)^n \left(\frac{1}{n+1}\right) \\
&= 0
\end{aligned}$$

Note that this limit is not as easily evaluated as the limit obtained by the Root Test in Example 4.

Strategies for Testing Series

You have now studied 10 tests for determining the convergence or divergence of an infinite series. (See the summary in the table on page 644.) Skill in choosing and applying the various tests will come only with practice. Below is a set of guidelines for choosing an appropriate test.

Guidelines for Testing a Series for Convergence or Divergence

1. Does the nth term approach 0? If not, the series diverges.
2. Is the series one of the special types—geometric, p-series, telescoping, or alternating?
3. Can the Integral Test, the Root Test, or the Ratio Test be applied?
4. Can the series be compared favorably to one of the special types?

In some instances, more than one test is applicable. However, your objective should be to learn to choose the most efficient test.

EXAMPLE 5 Applying the Strategies for Testing Series

Determine the convergence or divergence of each series.

a. $\displaystyle\sum_{n=1}^{\infty} \frac{n+1}{3n+1}$ **b.** $\displaystyle\sum_{n=1}^{\infty} \left(\frac{\pi}{6}\right)^n$ **c.** $\displaystyle\sum_{n=1}^{\infty} ne^{-n^2}$

d. $\displaystyle\sum_{n=1}^{\infty} \frac{1}{3n+1}$ **e.** $\displaystyle\sum_{n=1}^{\infty} (-1)^n \frac{3}{4n+1}$ **f.** $\displaystyle\sum_{n=1}^{\infty} \frac{n!}{10^n}$

g. $\displaystyle\sum_{n=1}^{\infty} \left(\frac{n+1}{2n+1}\right)^n$

Solution

a. For this series, the limit of the nth term is not 0 $\left(a_n \to \frac{1}{3} \text{ as } n \to \infty\right)$. So, by the nth-Term Test, the series diverges.

b. This series is geometric. Moreover, because the ratio $r = \pi/6$ of the terms is less than 1 in absolute value, you can conclude that the series converges.

c. Because the function $f(x) = xe^{-x^2}$ is easily integrated, you can use the Integral Test to conclude that the series converges.

d. The nth term of this series can be compared to the nth term of the harmonic series. After using the Limit Comparison Test, you can conclude that the series diverges.

e. This is an alternating series whose nth term approaches 0. Because $a_{n+1} \le a_n$, you can use the Alternating Series Test to conclude that the series converges.

f. The nth term of this series involves a factorial, which indicates that the Ratio Test may work well. After applying the Ratio Test, you can conclude that the series diverges.

g. The nth term of this series involves a variable that is raised to the nth power, which indicates that the Root Test may work well. After applying the Root Test, you can conclude that the series converges.

Try It Exploration A Exploration B

Summary of Tests for Series

Test	Series	Condition(s) of Convergence	Condition(s) of Divergence	Comment
nth-Term	$\sum_{n=1}^{\infty} a_n$		$\lim_{n\to\infty} a_n \neq 0$	This test cannot be used to show convergence.
Geometric Series	$\sum_{n=0}^{\infty} ar^n$	$\lvert r\rvert < 1$	$\lvert r\rvert \geq 1$	Sum: $S = \dfrac{a}{1-r}$
Telescoping Series	$\sum_{n=1}^{\infty} (b_n - b_{n+1})$	$\lim_{n\to\infty} b_n = L$		Sum: $S = b_1 - L$
p-Series	$\sum_{n=1}^{\infty} \dfrac{1}{n^p}$	$p > 1$	$p \leq 1$	
Alternating Series	$\sum_{n=1}^{\infty} (-1)^{n-1} a_n$	$0 < a_{n+1} \leq a_n$ and $\lim_{n\to\infty} a_n = 0$		Remainder: $\lvert R_N\rvert \leq a_{N+1}$
Integral (f is continuous, positive, and decreasing)	$\sum_{n=1}^{\infty} a_n$, $a_n = f(n) \geq 0$	$\int_1^{\infty} f(x)\, dx$ converges	$\int_1^{\infty} f(x)\, dx$ diverges	Remainder: $0 < R_N < \int_N^{\infty} f(x)\, dx$
Root	$\sum_{n=1}^{\infty} a_n$	$\lim_{n\to\infty} \sqrt[n]{\lvert a_n\rvert} < 1$	$\lim_{n\to\infty} \sqrt[n]{\lvert a_n\rvert} > 1$	Test is inconclusive if $\lim_{n\to\infty} \sqrt[n]{\lvert a_n\rvert} = 1$.
Ratio	$\sum_{n=1}^{\infty} a_n$	$\lim_{n\to\infty} \left\lvert \dfrac{a_{n+1}}{a_n} \right\rvert < 1$	$\lim_{n\to\infty} \left\lvert \dfrac{a_{n+1}}{a_n} \right\rvert > 1$	Test is inconclusive if $\lim_{n\to\infty} \left\lvert \dfrac{a_{n+1}}{a_n} \right\rvert = 1$.
Direct Comparison ($a_n, b_n > 0$)	$\sum_{n=1}^{\infty} a_n$	$0 < a_n \leq b_n$ and $\sum_{n=1}^{\infty} b_n$ converges	$0 < b_n \leq a_n$ and $\sum_{n=1}^{\infty} b_n$ diverges	
Limit Comparison ($a_n, b_n > 0$)	$\sum_{n=1}^{\infty} a_n$	$\lim_{n\to\infty} \dfrac{a_n}{b_n} = L > 0$ and $\sum_{n=1}^{\infty} b_n$ converges	$\lim_{n\to\infty} \dfrac{a_n}{b_n} = L > 0$ and $\sum_{n=1}^{\infty} b_n$ diverges	

Section 9.7 Taylor Polynomials and Approximations

- **Find polynomial approximations of elementary functions and compare them with the elementary functions.**
- **Find Taylor and Maclaurin polynomial approximations of elementary functions.**
- **Use the remainder of a Taylor polynomial.**

Polynomial Approximations of Elementary Functions

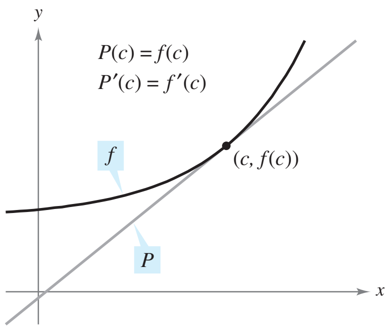
The goal of this section is to show how polynomial functions can be used as approximations for other elementary functions. To find a polynomial function P that approximates another function f, begin by choosing a number c in the domain of f at which f and P have the same value. That is,

$$P(c) = f(c). \qquad \text{Graphs of } f \text{ and } P \text{ pass through } (c, f(c)).$$

The approximating polynomial is said to be **expanded about c** or **centered at c.** Geometrically, the requirement that $P(c) = f(c)$ means that the graph of P passes through the point $(c, f(c))$. Of course, there are many polynomials whose graphs pass through the point $(c, f(c))$. Your task is to find a polynomial whose graph resembles the graph of f near this point. One way to do this is to impose the additional requirement that the slope of the polynomial function be the same as the slope of the graph of f at the point $(c, f(c))$.

$$P'(c) = f'(c) \qquad \text{Graphs of } f \text{ and } P \text{ have the same slope at } (c, f(c)).$$

With these two requirements, you can obtain a simple linear approximation of f, as shown in Figure 9.10.



Near $(c, f(c))$, the graph of P can be used to approximate the graph of f.
Figure 9.10

EXAMPLE 1 First-Degree Polynomial Approximation of $f(x) = e^x$

For the function $f(x) = e^x$, find a first-degree polynomial function

$$P_1(x) = a_0 + a_1x$$

whose value and slope agree with the value and slope of f at $x = 0$.

Solution Because $f(x) = e^x$ and $f'(x) = e^x$, the value and the slope of f, at $x = 0$, are given by

$$f(0) = e^0 = 1$$

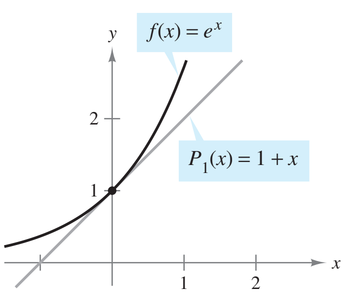
and

$$f'(0) = e^0 = 1.$$

Because $P_1(x) = a_0 + a_1x$, you can use the condition that $P_1(0) = f(0)$ to conclude that $a_0 = 1$. Moreover, because $P_1'(x) = a_1$, you can use the condition that $P_1'(0) = f'(0)$ to conclude that $a_1 = 1$. Therefore,

$$P_1(x) = 1 + x.$$

Figure 9.11 shows the graphs of $P_1(x) = 1 + x$ and $f(x) = e^x$.



P_1 is the first-degree polynomial approximation of $f(x) = e^x$.
Figure 9.11

Editable Graph

Try It Exploration A Exploration B

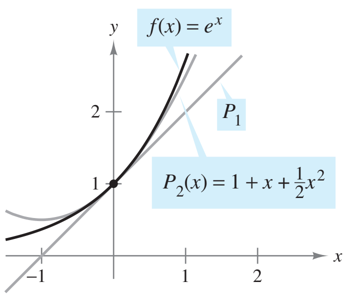
NOTE Example 1 isn't the first time you have used a linear function to approximate another function. The same procedure was used as the basis for Newton's Method.

In Figure 9.12 you can see that, at points near (0, 1), the graph of

$$P_1(x) = 1 + x \qquad \text{1st-degree approximation}$$

is reasonably close to the graph of $f(x) = e^x$. However, as you move away from (0, 1), the graphs move farther from each other and the accuracy of the approximation decreases. To improve the approximation, you can impose yet another requirement—that the values of the second derivatives of P and f agree when $x = 0$. The polynomial, P_2, of least degree that satisfies all three requirements $P_2(0) = f(0)$, $P_2'(0) = f'(0)$, and $P_2''(0) = f''(0)$ can be shown to be

$$P_2(x) = 1 + x + \frac{1}{2}x^2. \qquad \text{2nd-degree approximation}$$



P_2 is the second-degree polynomial approximation of $f(x) = e^x$.
Figure 9.12

Moreover, in Figure 9.12, you can see that P_2 is a better approximation of f than P_1. If you continue this pattern, requiring that the values of $P_n(x)$ and its first n derivatives match those of $f(x) = e^x$ at $x = 0$, you obtain the following.

$$\begin{aligned} P_n(x) &= 1 + x + \frac{1}{2}x^2 + \frac{1}{3!}x^3 + \cdot \cdot \cdot + \frac{1}{n!}x^n \qquad n\text{th-degree approximation} \\ &\approx e^x \end{aligned}$$

EXAMPLE 2 Third-Degree Polynomial Approximation of $f(x) = e^x$

Construct a table comparing the values of the polynomial

$$P_3(x) = 1 + x + \frac{1}{2}x^2 + \frac{1}{3!}x^3 \qquad \text{3rd-degree approximation}$$

with $f(x) = e^x$ for several values of x near 0.

Solution Using a calculator or a computer, you can obtain the results shown in the table. Note that for $x = 0$, the two functions have the same value, but that as x moves farther away from 0, the accuracy of the approximating polynomial $P_3(x)$ decreases.

x	−1.0	−0.2	−0.1	0	0.1	0.2	1.0
e^x	0.3679	0.81873	0.904837	1	1.105171	1.22140	2.7183
$P_3(x)$	0.3333	0.81867	0.904833	1	1.105167	1.22133	2.6667

Try It Exploration A

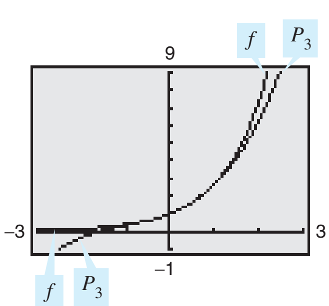
TECHNOLOGY A graphing utility can be used to compare the graph of the approximating polynomial with the graph of the function f. For instance, in Figure 9.13, the graph of

$$P_3(x) = 1 + x + \tfrac{1}{2}x^2 + \tfrac{1}{6}x^3 \qquad \text{3rd-degree approximation}$$

is compared with the graph of $f(x) = e^x$. If you have access to a graphing utility, try comparing the graphs of

$$P_4(x) = 1 + x + \tfrac{1}{2}x^2 + \tfrac{1}{6}x^3 + \tfrac{1}{24}x^4 \qquad \text{4th-degree approximation}$$
$$P_5(x) = 1 + x + \tfrac{1}{2}x^2 + \tfrac{1}{6}x^3 + \tfrac{1}{24}x^4 + \tfrac{1}{120}x^5 \qquad \text{5th-degree approximation}$$
$$P_6(x) = 1 + x + \tfrac{1}{2}x^2 + \tfrac{1}{6}x^3 + \tfrac{1}{24}x^4 + \tfrac{1}{120}x^5 + \tfrac{1}{720}x^6 \qquad \text{6th-degree approximation}$$

with the graph of f. What do you notice?



P_3 is the third-degree polynomial approximation of $f(x) = e^x$.
Figure 9.13

BROOK TAYLOR (1685–1731)

Although Taylor was not the first to seek polynomial approximations of transcendental functions, his account published in 1715 was one of the first comprehensive works on the subject.

MathBio

Taylor and Maclaurin Polynomials

The polynomial approximation of $f(x) = e^x$ given in Example 2 is expanded about $c = 0$. For expansions about an arbitrary value of c, it is convenient to write the polynomial in the form

$$P_n(x) = a_0 + a_1(x-c) + a_2(x-c)^2 + a_3(x-c)^3 + \cdots + a_n(x-c)^n.$$

In this form, repeated differentiation produces

$$\begin{aligned}
P_n'(x) &= a_1 + 2a_2(x-c) + 3a_3(x-c)^2 + \cdots + na_n(x-c)^{n-1}\\
P_n''(x) &= 2a_2 + 2(3a_3)(x-c) + \cdots + n(n-1)a_n(x-c)^{n-2}\\
P_n'''(x) &= 2(3a_3) + \cdots + n(n-1)(n-2)a_n(x-c)^{n-3}\\
&\vdots\\
P_n^{(n)}(x) &= n(n-1)(n-2)\cdots(2)(1)a_n.
\end{aligned}$$

Letting $x = c$, you then obtain

$$P_n(c) = a_0, \quad P_n'(c) = a_1, \quad P_n''(c) = 2a_2, \quad \ldots, \quad P_n^{(n)}(c) = n!a_n$$

and because the value of f and its first n derivatives must agree with the value of P_n and its first n derivatives at $x = c$, it follows that

$$f(c) = a_0, \quad f'(c) = a_1, \quad \frac{f''(c)}{2!} = a_2, \quad \ldots, \quad \frac{f^{(n)}(c)}{n!} = a_n.$$

With these coefficients, you can obtain the following definition of **Taylor polynomials,** named after the English mathematician Brook Taylor, and **Maclaurin polynomials,** named after the English mathematician Colin Maclaurin (1698–1746).

Definitions of *n*th Taylor Polynomial and *n*th Maclaurin Polynomial

If f has n derivatives at c, then the polynomial

$$P_n(x) = f(c) + f'(c)(x-c) + \frac{f''(c)}{2!}(x-c)^2 + \cdots + \frac{f^{(n)}(c)}{n!}(x-c)^n$$

is called the ***n*th Taylor polynomial for *f* at *c*.** If $c = 0$, then

$$P_n(x) = f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \frac{f'''(0)}{3!}x^3 + \cdots + \frac{f^{(n)}(0)}{n!}x^n$$

is also called the ***n*th Maclaurin polynomial for *f*.**

NOTE Maclaurin polynomials are special types of Taylor polynomials for which $c = 0$.

EXAMPLE 3 A Maclaurin Polynomial for $f(x) = e^x$

Find the nth Maclaurin polynomial for $f(x) = e^x$.

Solution From the discussion on page 649, the nth Maclaurin polynomial for

$$f(x) = e^x$$

is given by

$$P_n(x) = 1 + x + \frac{1}{2!}x^2 + \frac{1}{3!}x^3 + \cdots + \frac{1}{n!}x^n.$$

Select the animation button to see that the approximation becomes better as the degree of the Maclaurin polynomial increases.

Animation

FOR FURTHER INFORMATION To see how to use series to obtain other approximations to e, see the article "Novel Series-based Approximations to ***e***" by John Knox and Harlan J. Brothers in *The College Mathematics Journal*.

MathArticle

Try It Exploration A

EXAMPLE 4 Finding Taylor Polynomials for ln x

Find the Taylor polynomials P_0, P_1, P_2, P_3, and P_4 for $f(x) = \ln x$ centered at $c = 1$.

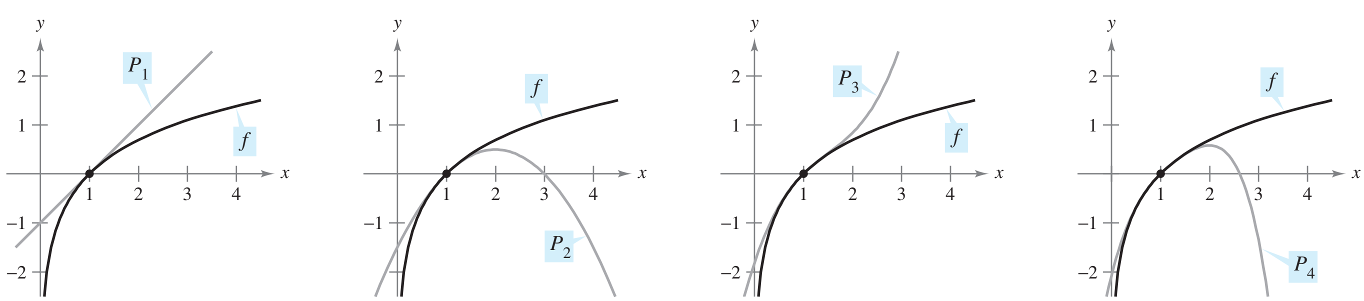
Solution Expanding about $c = 1$ yields the following.

$$\begin{aligned}
f(x) &= \ln x & f(1) &= \ln 1 = 0\\
f'(x) &= \frac{1}{x} & f'(1) &= \frac{1}{1} = 1\\
f''(x) &= -\frac{1}{x^2} & f''(1) &= -\frac{1}{1^2} = -1\\
f'''(x) &= \frac{2!}{x^3} & f'''(1) &= \frac{2!}{1^3} = 2\\
f^{(4)}(x) &= -\frac{3!}{x^4} & f^{(4)}(1) &= -\frac{3!}{1^4} = -6
\end{aligned}$$

Therefore, the Taylor polynomials are as follows.

$$\begin{aligned}
P_0(x) &= f(1) = 0\\
P_1(x) &= f(1) + f'(1)(x-1) = (x-1)\\
P_2(x) &= f(1) + f'(1)(x-1) + \frac{f''(1)}{2!}(x-1)^2\\
&= (x-1) - \frac{1}{2}(x-1)^2\\
P_3(x) &= f(1) + f'(1)(x-1) + \frac{f''(1)}{2!}(x-1)^2 + \frac{f'''(1)}{3!}(x-1)^3\\
&= (x-1) - \frac{1}{2}(x-1)^2 + \frac{1}{3}(x-1)^3\\
P_4(x) &= f(1) + f'(1)(x-1) + \frac{f''(1)}{2!}(x-1)^2 + \frac{f'''(1)}{3!}(x-1)^3\\
&\quad + \frac{f^{(4)}(1)}{4!}(x-1)^4\\
&= (x-1) - \frac{1}{2}(x-1)^2 + \frac{1}{3}(x-1)^3 - \frac{1}{4}(x-1)^4
\end{aligned}$$

Figure 9.14 compares the graphs of P_1, P_2, P_3, and P_4 with the graph of $f(x) = \ln x$. Note that near $x = 1$ the graphs are nearly indistinguishable. For instance, $P_4(0.9) \approx -0.105358$ and $\ln(0.9) \approx -0.105361$.



As n increases, the graph of P_n becomes a better and better approximation of the graph of $f(x) = \ln x$ near $x = 1$.
Figure 9.14

Try It Exploration A Exploration B

EXAMPLE 5 Finding Maclaurin Polynomials for cos x

Find the Maclaurin polynomials P_0, P_2, P_4, and P_6 for $f(x) = \cos x$. Use $P_6(x)$ to approximate the value of $\cos(0.1)$.

Solution Expanding about $c = 0$ yields the following.

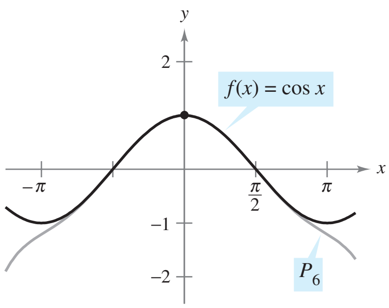
$$\begin{aligned} f(x) &= \cos x & f(0) &= \cos 0 = 1 \\ f'(x) &= -\sin x & f'(0) &= -\sin 0 = 0 \\ f''(x) &= -\cos x & f''(0) &= -\cos 0 = -1 \\ f'''(x) &= \sin x & f'''(0) &= \sin 0 = 0 \end{aligned}$$

Through repeated differentiation, you can see that the pattern $1, 0, -1, 0$ continues, and you obtain the following Maclaurin polynomials.

$$P_0(x) = 1, \quad P_2(x) = 1 - \frac{1}{2!}x^2,$$

$$P_4(x) = 1 - \frac{1}{2!}x^2 + \frac{1}{4!}x^4, \quad P_6(x) = 1 - \frac{1}{2!}x^2 + \frac{1}{4!}x^4 - \frac{1}{6!}x^6$$

Using $P_6(x)$, you obtain the approximation $\cos(0.1) \approx 0.995004165$, which coincides with the calculator value to nine decimal places. Figure 9.15 compares the graphs of $f(x) = \cos x$ and P_6.



Near $(0, 1)$, the graph of P_6 can be used to approximate the graph of $f(x) = \cos x$.
Figure 9.15

Editable Graph

Try It Exploration A

Note in Example 5 that the Maclaurin polynomials for $\cos x$ have only even powers of x. Similarly, the Maclaurin polynomials for $\sin x$ have only odd powers of x (see Exercise 17). This is not generally true of the Taylor polynomials for $\sin x$ and $\cos x$ expanded about $c \neq 0$, as you can see in the next example.

EXAMPLE 6 Finding a Taylor Polynomial for sin x

Find the third Taylor polynomial for $f(x) = \sin x$, expanded about $c = \pi/6$.

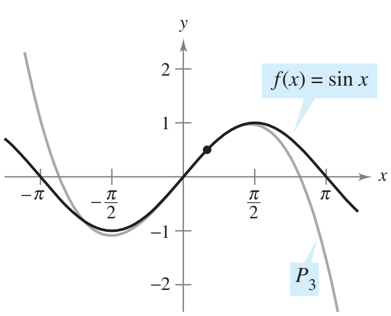
Solution Expanding about $c = \pi/6$ yields the following.

$$\begin{aligned} f(x) &= \sin x & f\left(\frac{\pi}{6}\right) &= \sin \frac{\pi}{6} = \frac{1}{2} \\ f'(x) &= \cos x & f'\left(\frac{\pi}{6}\right) &= \cos \frac{\pi}{6} = \frac{\sqrt{3}}{2} \\ f''(x) &= -\sin x & f''\left(\frac{\pi}{6}\right) &= -\sin \frac{\pi}{6} = -\frac{1}{2} \\ f'''(x) &= -\cos x & f'''\left(\frac{\pi}{6}\right) &= -\cos \frac{\pi}{6} = -\frac{\sqrt{3}}{2} \end{aligned}$$

So, the third Taylor polynomial for $f(x) = \sin x$, expanded about $c = \pi/6$, is

$$\begin{aligned} P_3(x) &= f\left(\frac{\pi}{6}\right) + f'\left(\frac{\pi}{6}\right)\left(x - \frac{\pi}{6}\right) + \frac{f''\left(\frac{\pi}{6}\right)}{2!}\left(x - \frac{\pi}{6}\right)^2 + \frac{f'''\left(\frac{\pi}{6}\right)}{3!}\left(x - \frac{\pi}{6}\right)^3 \\ &= \frac{1}{2} + \frac{\sqrt{3}}{2}\left(x - \frac{\pi}{6}\right) - \frac{1}{2(2!)}\left(x - \frac{\pi}{6}\right)^2 - \frac{\sqrt{3}}{2(3!)}\left(x - \frac{\pi}{6}\right)^3. \end{aligned}$$

Figure 9.16 compares the graphs of $f(x) = \sin x$ and P_3.



Near $(\pi/6, 1/2)$, the graph of P_3 can be used to approximate the graph of $f(x) = \sin x$.
Figure 9.16

Editable Graph

Try It Exploration A Open Exploration

Taylor polynomials and Maclaurin polynomials can be used to approximate the value of a function at a specific point. For instance, to approximate the value of $\ln(1.1)$, you can use Taylor polynomials for $f(x) = \ln x$ expanded about $c = 1$, as shown in Example 4, or you can use Maclaurin polynomials, as shown in Example 7.

EXAMPLE 7 Approximation Using Maclaurin Polynomials

Use a fourth Maclaurin polynomial to approximate the value of $\ln(1.1)$.

Solution Because 1.1 is closer to 1 than to 0, you should consider Maclaurin polynomials for the function $g(x) = \ln(1 + x)$.

$$\begin{aligned}
g(x) &= \ln(1+x) & g(0) &= \ln(1+0) = 0\\
g'(x) &= (1+x)^{-1} & g'(0) &= (1+0)^{-1} = 1\\
g''(x) &= -(1+x)^{-2} & g''(0) &= -(1+0)^{-2} = -1\\
g'''(x) &= 2(1+x)^{-3} & g'''(0) &= 2(1+0)^{-3} = 2\\
g^{(4)}(x) &= -6(1+x)^{-4} & g^{(4)}(0) &= -6(1+0)^{-4} = -6
\end{aligned}$$

Note that you obtain the same coefficients as in Example 4. Therefore, the fourth Maclaurin polynomial for $g(x) = \ln(1 + x)$ is

$$\begin{aligned}
P_4(x) &= g(0) + g'(0)x + \frac{g''(0)}{2!}x^2 + \frac{g'''(0)}{3!}x^3 + \frac{g^{(4)}(0)}{4!}x^4\\
&= x - \frac{1}{2}x^2 + \frac{1}{3}x^3 - \frac{1}{4}x^4.
\end{aligned}$$

Consequently,

$$\ln(1.1) = \ln(1 + 0.1) \approx P_4(0.1) \approx 0.0953083.$$

Check to see that the fourth Taylor polynomial (from Example 4), evaluated at $x = 1.1$, yields the same result.

Try It

Exploration A

n	$P_n(0.1)$
1	0.1000000
2	0.0950000
3	0.0953333
4	0.0953083

The table at the left illustrates the accuracy of the Taylor polynomial approximation of the calculator value of $\ln(1.1)$. You can see that as n becomes larger, $P_n(0.1)$ approaches the calculator value of 0.0953102.

On the other hand, the table below illustrates that as you move away from the expansion point $c = 1$, the accuracy of the approximation decreases.

Fourth Taylor Polynomial Approximation of $\ln(1 + x)$

x	0	0.1	0.5	0.75	1.0
$\ln(1+x)$	0	0.0953102	0.4054651	0.5596158	0.6931472
$P_4(x)$	0	0.0953083	0.4010417	0.5302734	0.5833333

These two tables illustrate two very important points about the accuracy of Taylor (or Maclaurin) polynomials for use in approximations.

1. The approximation is usually better at x-values close to c than at x-values far from c.
2. The approximation is usually better for higher-degree Taylor (or Maclaurin) polynomials than for those of lower degree.

Remainder of a Taylor Polynomial

An approximation technique is of little value without some idea of its accuracy. To measure the accuracy of approximating a function value $f(x)$ by the Taylor polynomial $P_n(x)$, you can use the concept of a **remainder** $R_n(x)$, defined as follows.

$$\underbrace{f(x)}_{\text{Exact value}} = \underbrace{P_n(x)}_{\text{Approximate Value}} + \underbrace{R_n(x)}_{\text{Remainder}}$$

So, $R_n(x) = f(x) - P_n(x)$. The absolute value of $R_n(x)$ is called the **error** associated with the approximation. That is,

$$\text{Error} = |R_n(x)| = |f(x) - P_n(x)|.$$

The next theorem gives a general procedure for estimating the remainder associated with a Taylor polynomial. This important theorem is called **Taylor's Theorem,** and the remainder given in the theorem is called the **Lagrange form of the remainder.** (The proof of the theorem is lengthy, and is given in Appendix A.)

THEOREM 9.19 Taylor's Theorem

If a function f is differentiable through order $n + 1$ in an interval I containing c, then, for each x in I, there exists z between x and c such that

$$f(x) = f(c) + f'(c)(x - c) + \frac{f''(c)}{2!}(x - c)^2 + \cdot \cdot \cdot + \frac{f^{(n)}(c)}{n!}(x - c)^n + R_n(x)$$

where

$$R_n(x) = \frac{f^{(n+1)}(z)}{(n + 1)!}(x - c)^{n+1}.$$

NOTE One useful consequence of Taylor's Theorem is that

$$|R_n(x)| \le \frac{|x - c|^{n+1}}{(n + 1)!} \max |f^{(n+1)}(z)|$$

where $\max|f^{(n+1)}(z)|$ is the maximum value of $f^{(n+1)}(z)$ between x and c.

For $n = 0$, Taylor's Theorem states that if f is differentiable in an interval I containing c, then, for each x in I, there exists z between x and c such that

$$f(x) = f(c) + f'(z)(x - c) \qquad \text{or} \qquad f'(z) = \frac{f(x) - f(c)}{x - c}.$$

Do you recognize this special case of Taylor's Theorem? (It is the Mean Value Theorem.)

When applying Taylor's Theorem, you should not expect to be able to find the exact value of z. (If you could do this, an approximation would not be necessary.) Rather, you try to find bounds for $f^{(n+1)}(z)$ from which you are able to tell how large the remainder $R_n(x)$ is.

EXAMPLE 8 Determining the Accuracy of an Approximation

The third Maclaurin polynomial for $\sin x$ is given by

$$P_3(x) = x - \frac{x^3}{3!}.$$

Use Taylor's Theorem to approximate $\sin(0.1)$ by $P_3(0.1)$ and determine the accuracy of the approximation.

Solution Using Taylor's Theorem, you have

$$\sin x = x - \frac{x^3}{3!} + R_3(x) = x - \frac{x^3}{3!} + \frac{f^{(4)}(z)}{4!}x^4$$

where $0 < z < 0.1$. Therefore,

$$\sin(0.1) \approx 0.1 - \frac{(0.1)^3}{3!} \approx 0.1 - 0.000167 = 0.099833.$$

NOTE Try using a calculator to verify the results obtained in Examples 8 and 9. For Example 8, you obtain

$$\sin(0.1) \approx 0.0998334.$$

For Example 9, you obtain

$$P_3(1.2) \approx 0.1827$$

and

$$\ln(1.2) \approx 0.1823.$$

Because $f^{(4)}(z) = \sin z$, it follows that the error $|R_3(0.1)|$ can be bounded as follows.

$$0 < R_3(0.1) = \frac{\sin z}{4!}(0.1)^4 < \frac{0.0001}{4!} \approx 0.000004$$

This implies that

$$0.099833 < \sin(0.1) = 0.099833 + R_3(x) < 0.099833 + 0.000004$$
$$0.099833 < \sin(0.1) < 0.099837.$$

Try It Exploration A

EXAMPLE 9 Approximating a Value to a Desired Accuracy

Determine the degree of the Taylor polynomial $P_n(x)$ expanded about $c = 1$ that should be used to approximate $\ln(1.2)$ so that the error is less than 0.001.

Solution Following the pattern of Example 4, you can see that the $(n + 1)$st derivative of $f(x) = \ln x$ is given by

$$f^{(n+1)}(x) = (-1)^n \frac{n!}{x^{n+1}}.$$

Using Taylor's Theorem, you know that the error $|R_n(1.2)|$ is given by

$$|R_n(1.2)| = \left|\frac{f^{(n+1)}(z)}{(n+1)!}(1.2 - 1)^{n+1}\right| = \frac{n!}{z^{n+1}}\left[\frac{1}{(n+1)!}\right](0.2)^{n+1}$$
$$= \frac{(0.2)^{n+1}}{z^{n+1}(n+1)}$$

where $1 < z < 1.2$. In this interval, $(0.2)^{n+1}/[z^{n+1}(n + 1)]$ is less than $(0.2)^{n+1}/(n + 1)$. So, you are seeking a value of n such that

$$\frac{(0.2)^{n+1}}{(n+1)} < 0.001 \quad \Longrightarrow \quad 1000 < (n + 1)5^{n+1}.$$

By trial and error, you can determine that the smallest value of n that satisfies this inequality is $n = 3$. So, you would need the third Taylor polynomial to achieve the desired accuracy in approximating $\ln(1.2)$.

Try It Exploration A

Section 9.8 Power Series

- **Understand the definition of a power series.**
- **Find the radius and interval of convergence of a power series.**
- **Determine the endpoint convergence of a power series.**
- **Differentiate and integrate a power series.**

EXPLORATION

Graphical Reasoning Use a graphing utility to approximate the graph of each power series near $x = 0$. (Use the first several terms of each series.) Each series represents a well-known function. What is the function?

a. $\displaystyle\sum_{n=0}^{\infty} \frac{(-1)^n x^n}{n!}$

b. $\displaystyle\sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!}$

c. $\displaystyle\sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)!}$

d. $\displaystyle\sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{2n+1}$

e. $\displaystyle\sum_{n=0}^{\infty} \frac{2^n x^n}{n!}$

Power Series

In Section 9.7, you were introduced to the concept of approximating functions by Taylor polynomials. For instance, the function $f(x) = e^x$ can be *approximated* by its Maclaurin polynomials as follows.

$$e^x \approx 1 + x \qquad \text{1st-degree polynomial}$$

$$e^x \approx 1 + x + \frac{x^2}{2!} \qquad \text{2nd-degree polynomial}$$

$$e^x \approx 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} \qquad \text{3rd-degree polynomial}$$

$$e^x \approx 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} \qquad \text{4th-degree polynomial}$$

$$e^x \approx 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^5}{5!} \qquad \text{5th-degree polynomial}$$

In that section, you saw that the higher the degree of the approximating polynomial, the better the approximation becomes.

In this and the next two sections, you will see that several important types of functions, including

$$f(x) = e^x$$

can be represented *exactly* by an infinite series called a **power series.** For example, the power series representation for e^x is

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdot\cdot\cdot + \frac{x^n}{n!} + \cdot\cdot\cdot .$$

For each real number x, it can be shown that the infinite series on the right converges to the number e^x. Before doing this, however, some preliminary results dealing with power series will be discussed—beginning with the following definition.

Definition of Power Series

If x is a variable, then an infinite series of the form

$$\sum_{n=0}^{\infty} a_n x^n = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \cdot\cdot\cdot + a_n x^n + \cdot\cdot\cdot$$

is called a **power series.** More generally, an infinite series of the form

$$\sum_{n=0}^{\infty} a_n (x - c)^n = a_0 + a_1(x - c) + a_2(x - c)^2 + \cdot\cdot\cdot + a_n(x - c)^n + \cdot\cdot\cdot$$

is called a **power series centered at *c*,** where c is a constant.

NOTE To simplify the notation for power series, we agree that $(x - c)^0 = 1$, even if $x = c$.

EXAMPLE 1 Power Series

a. The following power series is centered at 0.

$$\sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2} + \frac{x^3}{3!} + \cdots$$

b. The following power series is centered at -1.

$$\sum_{n=0}^{\infty} (-1)^n (x+1)^n = 1 - (x+1) + (x+1)^2 - (x+1)^3 + \cdots$$

c. The following power series is centered at 1.

$$\sum_{n=1}^{\infty} \frac{1}{n}(x-1)^n = (x-1) + \frac{1}{2}(x-1)^2 + \frac{1}{3}(x-1)^3 + \cdots$$

Try It Exploration A

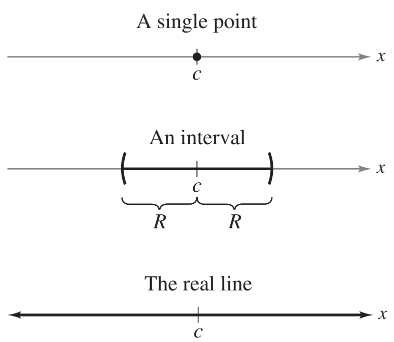
Radius and Interval of Convergence

A power series in x can be viewed as a function of x

$$f(x) = \sum_{n=0}^{\infty} a_n(x-c)^n$$

where the *domain of f* is the set of all x for which the power series converges. Determination of the domain of a power series is the primary concern in this section. Of course, every power series converges at its center c because

$$\begin{aligned} f(c) &= \sum_{n=0}^{\infty} a_n(c-c)^n \\ &= a_0(1) + 0 + 0 + \cdots + 0 + \cdots \\ &= a_0. \end{aligned}$$



The domain of a power series has only three basic forms: a single point, an interval centered at c, or the entire real line.
Figure 9.17

So, c always lies in the domain of f. The following important theorem states that the domain of a power series can take three basic forms: a single point, an interval centered at c, or the entire real line, as shown in Figure 9.17. A proof is given in Appendix A.

THEOREM 9.20 Convergence of a Power Series

For a power series centered at c, precisely one of the following is true.

1. The series converges only at c.
2. There exists a real number $R > 0$ such that the series converges absolutely for $|x - c| < R$, and diverges for $|x - c| > R$.
3. The series converges absolutely for all x.

The number R is the **radius of convergence** of the power series. If the series converges only at c, the radius of convergence is $R = 0$, and if the series converges for all x, the radius of convergence is $R = \infty$. The set of all values of x for which the power series converges is the **interval of convergence** of the power series.

STUDY TIP To determine the radius of convergence of a power series, use the Ratio Test, as demonstrated in Examples 2, 3, and 4.

EXAMPLE 2 Finding the Radius of Convergence

Find the radius of convergence of $\sum_{n=0}^{\infty} n!x^n$.

Solution For $x = 0$, you obtain

$$f(0) = \sum_{n=0}^{\infty} n!0^n = 1 + 0 + 0 + \cdot \cdot \cdot = 1.$$

For any fixed value of x such that $|x| > 0$, let $u_n = n!x^n$. Then

$$\begin{aligned}\lim_{n\to\infty}\left|\frac{u_{n+1}}{u_n}\right| &= \lim_{n\to\infty}\left|\frac{(n+1)!x^{n+1}}{n!x^n}\right| \\ &= |x| \lim_{n\to\infty}(n+1) \\ &= \infty.\end{aligned}$$

Therefore, by the Ratio Test, the series diverges for $|x| > 0$ and converges only at its center, 0. So, the radius of convergence is $R = 0$.

Try It | Exploration A | Exploration B

EXAMPLE 3 Finding the Radius of Convergence

Find the radius of convergence of

$$\sum_{n=0}^{\infty} 3(x-2)^n.$$

Solution For $x \neq 2$, let $u_n = 3(x-2)^n$. Then

$$\begin{aligned}\lim_{n\to\infty}\left|\frac{u_{n+1}}{u_n}\right| &= \lim_{n\to\infty}\left|\frac{3(x-2)^{n+1}}{3(x-2)^n}\right| \\ &= \lim_{n\to\infty}|x-2| \\ &= |x-2|.\end{aligned}$$

By the Ratio Test, the series converges if $|x-2| < 1$ and diverges if $|x-2| > 1$. Therefore, the radius of convergence of the series is $R = 1$.

Try It | Exploration A

EXAMPLE 4 Finding the Radius of Convergence

Find the radius of convergence of

$$\sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)!}.$$

Solution Let $u_n = (-1)^n x^{2n+1}/(2n+1)!$. Then

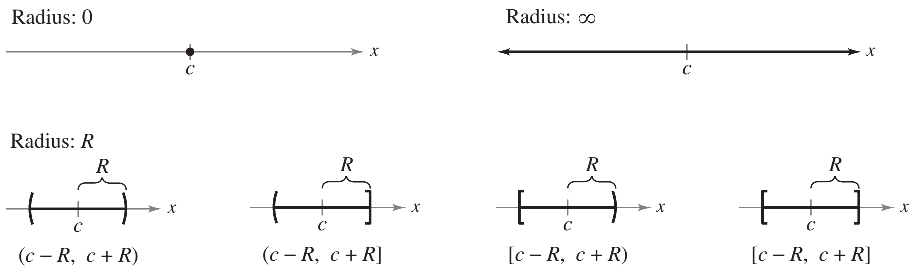
$$\begin{aligned}\lim_{n\to\infty}\left|\frac{u_{n+1}}{u_n}\right| &= \lim_{n\to\infty}\left|\frac{\dfrac{(-1)^{n+1}x^{2n+3}}{(2n+3)!}}{\dfrac{(-1)^n x^{2n+1}}{(2n-1)!}}\right| \\ &= \lim_{n\to\infty}\frac{x^2}{(2n+3)(2n+2)}.\end{aligned}$$

For any *fixed* value of x, this limit is 0. So, by the Ratio Test, the series converges for all x. Therefore, the radius of convergence is $R = \infty$.

Try It | Exploration A

Endpoint Convergence

Note that for a power series whose radius of convergence is a finite number R, Theorem 9.20 says nothing about the convergence at the *endpoints* of the interval of convergence. Each endpoint must be tested separately for convergence or divergence. As a result, the interval of convergence of a power series can take any one of the six forms shown in Figure 9.18.



Intervals of convergence
Figure 9.18

EXAMPLE 5 Finding the Interval of Convergence

Find the interval of convergence of $\displaystyle\sum_{n=1}^{\infty} \frac{x^n}{n}$.

Solution Letting $u_n = x^n/n$ produces

$$\lim_{n\to\infty}\left|\frac{u_{n+1}}{u_n}\right| = \lim_{n\to\infty}\left|\frac{\dfrac{x^{n+1}}{(n+1)}}{\dfrac{x^n}{n}}\right|$$
$$= \lim_{n\to\infty}\left|\frac{nx}{n+1}\right|$$
$$= |x|.$$

So, by the Ratio Test, the radius of convergence is $R = 1$. Moreover, because the series is centered at 0, it converges in the interval $(-1, 1)$. This interval, however, is not necessarily the *interval of convergence*. To determine this, you must test for convergence at each endpoint. When $x = 1$, you obtain the *divergent* harmonic series

$$\sum_{n=1}^{\infty}\frac{1}{n} = \frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \cdot\cdot\cdot. \qquad \text{Diverges when } x = 1$$

When $x = -1$, you obtain the *convergent* alternating harmonic series

$$\sum_{n=1}^{\infty}\frac{(-1)^n}{n} = -1 + \frac{1}{2} - \frac{1}{3} + \frac{1}{4} - \cdot\cdot\cdot. \qquad \text{Converges when } x = -1$$

So, the interval of convergence for the series is $[-1, 1)$, as shown in Figure 9.19.

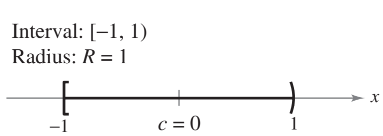


Figure 9.19

Try It Exploration A Open Exploration

EXAMPLE 6 Finding the Interval of Convergence

Find the interval of convergence of $\displaystyle\sum_{n=0}^{\infty} \frac{(-1)^n(x+1)^n}{2^n}$.

Solution Letting $u_n = (-1)^n(x+1)^n/2^n$ produces

$$\lim_{n\to\infty}\left|\frac{u_{n+1}}{u_n}\right| = \lim_{n\to\infty}\left|\frac{\dfrac{(-1)^{n+1}(x+1)^{n+1}}{2^{n+1}}}{\dfrac{(-1)^n(x+1)^n}{2^n}}\right|$$

$$= \lim_{n\to\infty}\left|\frac{2^n(x+1)}{2^{n+1}}\right|$$

$$= \left|\frac{x+1}{2}\right|.$$

By the Ratio Test, the series converges if $|(x+1)/2| < 1$ or $|x+1| < 2$. So, the radius of convergence is $R = 2$. Because the series is centered at $x = -1$, it will converge in the interval $(-3, 1)$. Furthermore, at the endpoints you have

$$\sum_{n=0}^{\infty} \frac{(-1)^n(-2)^n}{2^n} = \sum_{n=0}^{\infty} \frac{2^n}{2^n} = \sum_{n=0}^{\infty} 1 \qquad \text{Diverges when } x = -3$$

and

$$\sum_{n=0}^{\infty} \frac{(-1)^n(2)^n}{2^n} = \sum_{n=0}^{\infty} (-1)^n \qquad \text{Diverges when } x = 1$$

both of which diverge. So, the interval of convergence is $(-3, 1)$, as shown in Figure 9.20.

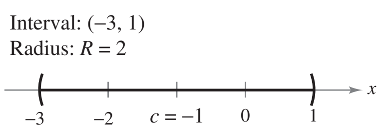


Figure 9.20

Try It   Exploration A

EXAMPLE 7 Finding the Interval of Convergence

Find the interval of convergence of

$\displaystyle\sum_{n=1}^{\infty} \frac{x^n}{n^2}$.

Solution Letting $u_n = x^n/n^2$ produces

$$\lim_{n\to\infty}\left|\frac{u_{n+1}}{u_n}\right| = \lim_{n\to\infty}\left|\frac{x^{n+1}/(n+1)^2}{x^n/n^2}\right|$$

$$= \lim_{n\to\infty}\left|\frac{n^2x}{(n+1)^2}\right| = |x|.$$

So, the radius of convergence is $R = 1$. Because the series is centered at $x = 0$, it converges in the interval $(-1, 1)$. When $x = 1$, you obtain the convergent p-series

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{1}{1^2} + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{4^2} + \cdots. \qquad \text{Converges when } x = 1$$

When $x = -1$, you obtain the convergent alternating series

$$\sum_{n=1}^{\infty} \frac{(-1)^n}{n^2} = -\frac{1}{1^2} + \frac{1}{2^2} - \frac{1}{3^2} + \frac{1}{4^2} - \cdots. \qquad \text{Converges when } x = -1$$

Therefore, the interval of convergence for the given series is $[-1, 1]$.

Try It   Exploration A

JAMES GREGORY (1638–1675)

One of the earliest mathematicians to work with power series was a Scotsman, James Gregory. He developed a power series method for interpolating table values—a method that was later used by Brook Taylor in the development of Taylor polynomials and Taylor series.

Differentiation and Integration of Power Series

Power series representation of functions has played an important role in the development of calculus. In fact, much of Newton's work with differentiation and integration was done in the context of power series—especially his work with complicated algebraic functions and transcendental functions. Euler, Lagrange, Leibniz, and the Bernoullis all used power series extensively in calculus.

Once you have defined a function with a power series, it is natural to wonder how you can determine the characteristics of the function. Is it continuous? Differentiable? Theorem 9.21, which is stated without proof, answers these questions.

THEOREM 9.21 Properties of Functions Defined by Power Series

If the function given by

$$\begin{aligned} f(x) &= \sum_{n=0}^{\infty} a_n(x-c)^n \\ &= a_0 + a_1(x-c) + a_2(x-c)^2 + a_3(x-c)^3 + \cdots \end{aligned}$$

has a radius of convergence of $R > 0$, then, on the interval $(c - R, c + R)$, f is differentiable (and therefore continuous). Moreover, the derivative and antiderivative of f are as follows.

1. $$\begin{aligned} f'(x) &= \sum_{n=1}^{\infty} na_n(x-c)^{n-1} \\ &= a_1 + 2a_2(x-c) + 3a_3(x-c)^2 + \cdots \end{aligned}$$

2. $$\begin{aligned} \int f(x)\,dx &= C + \sum_{n=0}^{\infty} a_n \frac{(x-c)^{n+1}}{n+1} \\ &= C + a_0(x-c) + a_1\frac{(x-c)^2}{2} + a_2\frac{(x-c)^3}{3} + \cdots \end{aligned}$$

The *radius of convergence* of the series obtained by differentiating or integrating a power series is the same as that of the original power series. The *interval of convergence*, however, may differ as a result of the behavior at the endpoints.

Theorem 9.21 states that, in many ways, a function defined by a power series behaves like a polynomial. It is continuous in its interval of convergence, and both its derivative and its antiderivative can be determined by differentiating and integrating each term of the given power series. For instance, the derivative of the power series

$$\begin{aligned} f(x) &= \sum_{n=0}^{\infty} \frac{x^n}{n!} \\ &= 1 + x + \frac{x^2}{2} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots \end{aligned}$$

is

$$\begin{aligned} f'(x) &= 1 + (2)\frac{x}{2} + (3)\frac{x^2}{3!} + (4)\frac{x^3}{4!} + \cdots \\ &= 1 + x + \frac{x^2}{2} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots \\ &= f(x). \end{aligned}$$

Notice that $f'(x) = f(x)$. Do you recognize this function?

EXAMPLE 8 Intervals of Convergence for $f(x)$, $f'(x)$, and $\int f(x)\,dx$

Consider the function given by

$$f(x) = \sum_{n=1}^{\infty} \frac{x^n}{n} = x + \frac{x^2}{2} + \frac{x^3}{3} + \cdot \cdot \cdot .$$

Find the intervals of convergence for each of the following.

a. $\int f(x)\,dx$ **b.** $f(x)$ **c.** $f'(x)$

Solution By Theorem 9.21, you have

$$\begin{aligned} f'(x) &= \sum_{n=1}^{\infty} x^{n-1} \\ &= 1 + x + x^2 + x^3 + \cdot \cdot \cdot \end{aligned}$$

and

$$\begin{aligned} \int f(x)\,dx &= C + \sum_{n=1}^{\infty} \frac{x^{n+1}}{n(n+1)} \\ &= C + \frac{x^2}{1 \cdot 2} + \frac{x^3}{2 \cdot 3} + \frac{x^4}{3 \cdot 4} + \cdot \cdot \cdot . \end{aligned}$$

By the Ratio Test, you can show that each series has a radius of convergence of $R = 1$. Considering the interval $(-1, 1)$, you have the following.

a. For $\int f(x)\,dx$, the series

$$\sum_{n=1}^{\infty} \frac{x^{n+1}}{n(n+1)} \qquad \text{Interval of convergence: } [-1, 1]$$

converges for $x = \pm 1$, and its interval of convergence is $[-1, 1]$. See Figure 9.21(a).

b. For $f(x)$, the series

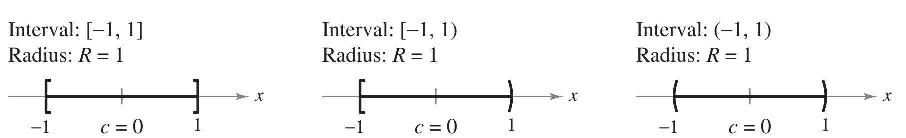
$$\sum_{n=1}^{\infty} \frac{x^n}{n} \qquad \text{Interval of convergence: } [-1, 1)$$

converges for $x = -1$ and diverges for $x = 1$. So, its interval of convergence is $[-1, 1)$. See Figure 9.21(b).

c. For $f'(x)$, the series

$$\sum_{n=1}^{\infty} x^{n-1} \qquad \text{Interval of convergence: } (-1, 1)$$

diverges for $x = \pm 1$, and its interval of convergence is $(-1, 1)$. See Figure 9.21(c).



(a) **(b)** **(c)**

Figure 9.21

Try It | Exploration A | Exploration B

From Example 8, it appears that of the three series, the one for the derivative, $f'(x)$, is the least likely to converge at the endpoints. In fact, it can be shown that if the series for $f'(x)$ converges at the endpoints $x = c \pm R$, the series for $f(x)$ will also converge there.

Section 9.9 Representation of Functions by Power Series

- **Find a geometric power series that represents a function.**
- **Construct a power series using series operations.**

> **JOSEPH FOURIER (1768–1830)**
>
> Some of the early work in representing functions by power series was done by the French mathematician Joseph Fourier. Fourier's work is important in the history of calculus, partly because it forced eighteenth century mathematicians to question the then-prevailing narrow concept of a function. Both Cauchy and Dirichlet were motivated by Fourier's work with series, and in 1837 Dirichlet published the general definition of a function that is used today.

Geometric Power Series

In this section and the next, you will study several techniques for finding a power series that represents a given function.

Consider the function given by $f(x) = 1/(1 - x)$. The form of f closely resembles the sum of a geometric series

$$\sum_{n=0}^{\infty} ar^n = \frac{a}{1 - r}, \quad |r| < 1.$$

In other words, if you let $a = 1$ and $r = x$, a power series representation for $1/(1 - x)$, centered at 0, is

$$\begin{aligned}\frac{1}{1 - x} &= \sum_{n=0}^{\infty} x^n \\ &= 1 + x + x^2 + x^3 + \cdot \cdot \cdot, \quad |x| < 1.\end{aligned}$$

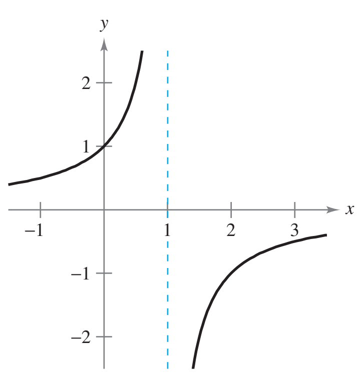
Of course, this series represents $f(x) = 1/(1 - x)$ only on the interval $(-1, 1)$, whereas f is defined for all $x \neq 1$, as shown in Figure 9.22. To represent f in another interval, you must develop a different series. For instance, to obtain the power series centered at -1, you could write

$$\frac{1}{1 - x} = \frac{1}{2 - (x + 1)} = \frac{1/2}{1 - [(x + 1)/2]} = \frac{a}{1 - r}$$

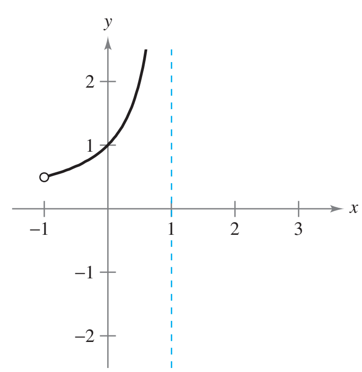
which implies that $a = \frac{1}{2}$ and $r = (x + 1)/2$. So, for $|x + 1| < 2$, you have

$$\begin{aligned}\frac{1}{1 - x} &= \sum_{n=0}^{\infty} \left(\frac{1}{2}\right)\left(\frac{x + 1}{2}\right)^n \\ &= \frac{1}{2}\left[1 + \frac{(x + 1)}{2} + \frac{(x + 1)^2}{4} + \frac{(x + 1)^3}{8} + \cdot \cdot \cdot\right], \quad |x + 1| < 2\end{aligned}$$

which converges on the interval $(-3, 1)$.



$f(x) = \dfrac{1}{1 - x}$, Domain: all $x \neq 1$



$f(x) = \sum_{n=0}^{\infty} x^n$, Domain: $-1 < x < 1$

Figure 9.22

EXAMPLE 1 Finding a Geometric Power Series Centered at 0

Find a power series for $f(x) = \dfrac{4}{x+2}$, centered at 0.

Solution Writing $f(x)$ in the form $a/(1-r)$ produces

$$\frac{4}{2+x} = \frac{2}{1-(-x/2)} = \frac{a}{1-r}$$

which implies that $a = 2$ and $r = -x/2$. So, the power series for $f(x)$ is

$$\begin{aligned}\frac{4}{x+2} &= \sum_{n=0}^{\infty} ar^n \\ &= \sum_{n=0}^{\infty} 2\left(-\frac{x}{2}\right)^n \\ &= 2\left(1 - \frac{x}{2} + \frac{x^2}{4} - \frac{x^3}{8} + \cdot \cdot \cdot\right).\end{aligned}$$

This power series converges when

$$\left|-\frac{x}{2}\right| < 1$$

which implies that the interval of convergence is $(-2, 2)$.

Try It | Exploration A | Technology

Long Division

$$\begin{array}{r l}
 & 2 - x + \frac{1}{2}x^2 - \frac{1}{4}x^3 + \cdot \cdot \cdot \\ \hline
2 + x \,\big) & 4 \\
 & \underline{4 + 2x} \\
 & -2x \\
 & \underline{-2x - x^2} \\
 & x^2 \\
 & \underline{x^2 + \frac{1}{2}x^3} \\
 & -\frac{1}{2}x^3 \\
 & \underline{-\frac{1}{2}x^3 - \frac{1}{4}x^4}
\end{array}$$

Another way to determine a power series for a rational function such as the one in Example 1 is to use long division. For instance, by dividing $2 + x$ into 4, you obtain the result shown at the left.

EXAMPLE 2 Finding a Geometric Power Series Centered at 1

Find a power series for $f(x) = \dfrac{1}{x}$, centered at 1.

Solution Writing $f(x)$ in the form $a/(1-r)$ produces

$$\frac{1}{x} = \frac{1}{1-(-x+1)} = \frac{a}{1-r}$$

which implies that $a = 1$ and $r = 1 - x = -(x-1)$. So, the power series for $f(x)$ is

$$\begin{aligned}\frac{1}{x} &= \sum_{n=0}^{\infty} ar^n \\ &= \sum_{n=0}^{\infty} [-(x-1)]^n \\ &= \sum_{n=0}^{\infty} (-1)^n(x-1)^n \\ &= 1 - (x-1) + (x-1)^2 - (x-1)^3 + \cdot \cdot \cdot .\end{aligned}$$

This power series converges when

$$|x-1| < 1$$

which implies that the interval of convergence is $(0, 2)$.

Try It | Exploration A

Operations with Power Series

The versatility of geometric power series will be shown later in this section, following a discussion of power series operations. These operations, used with differentiation and integration, provide a means of developing power series for a variety of elementary functions. (For simplicity, the following properties are stated for a series centered at 0.)

Operations with Power Series

Let $f(x) = \Sigma\, a_n x^n$ and $g(x) = \Sigma\, b_n x^n$.

1. $f(kx) = \displaystyle\sum_{n=0}^{\infty} a_n k^n x^n$

2. $f(x^N) = \displaystyle\sum_{n=0}^{\infty} a_n x^{nN}$

3. $f(x) \pm g(x) = \displaystyle\sum_{n=0}^{\infty} (a_n \pm b_n) x^n$

The operations described above can change the interval of convergence for the resulting series. For example, in the following addition, the interval of convergence for the sum is the *intersection* of the intervals of convergence of the two original series.

$$\underbrace{\sum_{n=0}^{\infty} x^n}_{(-1,\,1)} + \underbrace{\sum_{n=0}^{\infty} \left(\frac{x}{2}\right)^n}_{\cap\ (-2,\,2)} = \underbrace{\sum_{n=0}^{\infty} \left(1 + \frac{1}{2^n}\right) x^n}_{=\ (-1,\,1)}$$

EXAMPLE 3 Adding Two Power Series

Find a power series, centered at 0, for $f(x) = \dfrac{3x - 1}{x^2 - 1}$.

Solution Using partial fractions, you can write $f(x)$ as

$$\frac{3x - 1}{x^2 - 1} = \frac{2}{x + 1} + \frac{1}{x - 1}.$$

By adding the two geometric power series

$$\frac{2}{x + 1} = \frac{2}{1 - (-x)} = \sum_{n=0}^{\infty} 2(-1)^n x^n, \quad |x| < 1$$

and

$$\frac{1}{x - 1} = \frac{-1}{1 - x} = -\sum_{n=0}^{\infty} x^n, \quad |x| < 1$$

you obtain the following power series.

$$\frac{3x - 1}{x^2 - 1} = \sum_{n=0}^{\infty} [2(-1)^n - 1] x^n = 1 - 3x + x^2 - 3x^3 + x^4 - \cdots$$

The interval of convergence for this power series is $(-1, 1)$.

Try It **Exploration A**

EXAMPLE 4 Finding a Power Series by Integration

Find a power series for $f(x) = \ln x$, centered at 1.

Solution From Example 2, you know that

$$\frac{1}{x} = \sum_{n=0}^{\infty} (-1)^n (x-1)^n. \qquad \text{Interval of convergence: } (0, 2)$$

Integrating this series produces

$$\begin{aligned}\ln x &= \int \frac{1}{x}\,dx + C \\ &= C + \sum_{n=0}^{\infty} (-1)^n \frac{(x-1)^{n+1}}{n+1}.\end{aligned}$$

By letting $x = 1$, you can conclude that $C = 0$. Therefore,

$$\begin{aligned}\ln x &= \sum_{n=0}^{\infty} (-1)^n \frac{(x-1)^{n+1}}{n+1} \\ &= \frac{(x-1)}{1} - \frac{(x-1)^2}{2} + \frac{(x-1)^3}{3} - \frac{(x-1)^4}{4} + \cdots. \qquad \text{Interval of convergence: } (0, 2]\end{aligned}$$

Note that the series converges at $x = 2$. This is consistent with the observation in the preceding section that integration of a power series may alter the convergence at the endpoints of the interval of convergence.

Try It | Exploration A | Exploration B | Exploration C

TECHNOLOGY In Section 9.7, the fourth-degree Taylor polynomial for the natural logarithmic function

$$\ln x \approx (x-1) - \frac{(x-1)^2}{2} + \frac{(x-1)^3}{3} - \frac{(x-1)^4}{4}$$

was used to approximate $\ln(1.1)$.

$$\begin{aligned}\ln(1.1) &\approx (0.1) - \frac{1}{2}(0.1)^2 + \frac{1}{3}(0.1)^3 - \frac{1}{4}(0.1)^4 \\ &\approx 0.0953083\end{aligned}$$

You now know from Example 4 that this polynomial represents the first four terms of the power series for $\ln x$. Moreover, using the Alternating Series Remainder, you can determine that the error in this approximation is less than

$$\begin{aligned}|R_4| &\le |a_5| \\ &= \frac{1}{5}(0.1)^5 \\ &= 0.000002.\end{aligned}$$

During the seventeenth and eighteenth centuries, mathematical tables for logarithms and values of other transcendental functions were computed in this manner. Such numerical techniques are far from outdated, because it is precisely by such means that many modern calculating devices are programmed to evaluate transcendental functions.

EXAMPLE 5 Finding a Power Series by Integration

Find a power series for $g(x) = \arctan x$, centered at 0.

Solution Because $D_x[\arctan x] = 1/(1 + x^2)$, you can use the series

$$f(x) = \frac{1}{1+x} = \sum_{n=0}^{\infty} (-1)^n x^n. \qquad \text{Interval of convergence: } (-1, 1)$$

Substituting x^2 for x produces

$$f(x^2) = \frac{1}{1+x^2} = \sum_{n=0}^{\infty} (-1)^n x^{2n}.$$

Finally, by integrating, you obtain

$$\begin{aligned}
\arctan x &= \int \frac{1}{1+x^2}\,dx + C \\
&= C + \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{2n+1} \\
&= \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{2n+1} && \text{Let } x = 0, \text{ then } C = 0. \\
&= x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \cdots. && \text{Interval of convergence: } (-1, 1)
\end{aligned}$$

Try It | Exploration A | Open Exploration

It can be shown that the power series developed for $\arctan x$ in Example 5 also converges (to $\arctan x$) for $x = \pm 1$. For instance, when $x = 1$, you can write

$$\begin{aligned}
\arctan 1 &= 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots \\
&= \frac{\pi}{4}.
\end{aligned}$$

However, this series (developed by James Gregory in 1671) does not give us a practical way of approximating π because it converges so slowly that hundreds of terms would have to be used to obtain reasonable accuracy. Example 6 shows how to use *two* different arctangent series to obtain a very good approximation of π using only a few terms. This approximation was developed by John Machin in 1706.

SRINIVASA RAMANUJAN (1887–1920)

Series that can be used to approximate π have interested mathematicians for the past 300 years. An amazing series for approximating $1/\pi$ was discovered by the Indian mathematician Srinivasa Ramanujan in 1914 (see Exercise 64). Each successive term of Ramanujan's series adds roughly eight more correct digits to the value of $1/\pi$. For more information about Ramanujan's work, see the article "Ramanujan and Pi" by Jonathan M. Borwein and Peter B. Borwein in *Scientific American*.

MathBio

EXAMPLE 6 Approximating π with a Series

Use the trigonometric identity

$$4 \arctan \frac{1}{5} - \arctan \frac{1}{239} = \frac{\pi}{4}$$

to approximate the number π [see Exercise 50(b)].

Solution By using only five terms from each of the series for $\arctan(1/5)$ and $\arctan(1/239)$, you obtain

$$4\left(4 \arctan \frac{1}{5} - \arctan \frac{1}{239}\right) \approx 3.1415926$$

which agrees with the exact value of π with an error of less than 0.0000001.

Try It | Exploration A

Section 9.10 Taylor and Maclaurin Series

- **Find a Taylor or Maclaurin series for a function.**
- **Find a binomial series.**
- **Use a basic list of Taylor series to find other Taylor series.**

COLIN MACLAURIN (1698–1746)

The development of power series to represent functions is credited to the combined work of many seventeenth and eighteenth century mathematicians. Gregory, Newton, John and James Bernoulli, Euler, Lagrange, Wallis, and Fourier all contributed to this work. However, the two names that are most commonly associated with power series are Brook Taylor (1685–1731) and Colin Maclaurin.

MathBio

Taylor Series and Maclaurin Series

In Section 9.9, you derived power series for several functions using geometric series with term-by-term differentiation or integration. In this section you will study a *general* procedure for deriving the power series for a function that has derivatives of all orders. The following theorem gives the form that *every* convergent power series must take.

THEOREM 9.22 The Form of a Convergent Power Series

If f is represented by a power series $f(x) = \Sigma\, a_n(x - c)^n$ for all x in an open interval I containing c, then $a_n = f^{(n)}(c)/n!$ and

$$f(x) = f(c) + f'(c)(x - c) + \frac{f''(c)}{2!}(x - c)^2 + \cdots + \frac{f^{(n)}(c)}{n!}(x - c)^n + \cdots.$$

Proof Suppose the power series $\Sigma\, a_n(x - c)^n$ has a radius of convergence R. Then, by Theorem 9.21, you know that the nth derivative of f exists for $|x - c| < R$, and by successive differentiation you obtain the following.

$$\begin{aligned}
f^{(0)}(x) &= a_0 + a_1(x - c) + a_2(x - c)^2 + a_3(x - c)^3 + a_4(x - c)^4 + \cdots \\
f^{(1)}(x) &= a_1 + 2a_2(x - c) + 3a_3(x - c)^2 + 4a_4(x - c)^3 + \cdots \\
f^{(2)}(x) &= 2a_2 + 3!a_3(x - c) + 4 \cdot 3a_4(x - c)^2 + \cdots \\
f^{(3)}(x) &= 3!a_3 + 4!a_4(x - c) + \cdots \\
&\vdots \\
f^{(n)}(x) &= n!a_n + (n + 1)!a_{n+1}(x - c) + \cdots
\end{aligned}$$

Evaluating each of these derivatives at $x = c$ yields

$$\begin{aligned}
f^{(0)}(c) &= 0!a_0 \\
f^{(1)}(c) &= 1!a_1 \\
f^{(2)}(c) &= 2!a_2 \\
f^{(3)}(c) &= 3!a_3
\end{aligned}$$

and, in general, $f^{(n)}(c) = n!a_n$. By solving for a_n, you find that the coefficients of the power series representation of $f(x)$ are

$$a_n = \frac{f^{(n)}(c)}{n!}.$$

NOTE Be sure you understand Theorem 9.22. The theorem says that *if a power series converges to* $f(x)$, the series must be a Taylor series. The theorem does *not* say that every series formed with the Taylor coefficients $a_n = f^{(n)}(c)/n!$ will converge to $f(x)$.

Notice that the coefficients of the power series in Theorem 9.22 are precisely the coefficients of the Taylor polynomials for $f(x)$ at c as defined in Section 9.7. For this reason, the series is called the **Taylor series** for $f(x)$ at c.

Definitions of Taylor and Maclaurin Series

If a function f has derivatives of all orders at $x = c$, then the series

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(c)}{n!}(x-c)^n = f(c) + f'(c)(x-c) + \cdot \cdot \cdot + \frac{f^{(n)}(c)}{n!}(x-c)^n + \cdot \cdot \cdot$$

is called the **Taylor series for $f(x)$ at c.** Moreover, if $c = 0$, then the series is the **Maclaurin series for f.**

If you know the pattern for the coefficients of the Taylor polynomials for a function, you can extend the pattern easily to form the corresponding Taylor series. For instance, in Example 4 in Section 9.7, you found the fourth Taylor polynomial for $\ln x$, centered at 1, to be

$$P_4(x) = (x-1) - \frac{1}{2}(x-1)^2 + \frac{1}{3}(x-1)^3 - \frac{1}{4}(x-1)^4.$$

From this pattern, you can obtain the Taylor series for $\ln x$ centered at $c = 1$,

$$(x-1) - \frac{1}{2}(x-1)^2 + \cdot \cdot \cdot + \frac{(-1)^{n+1}}{n}(x-1)^n + \cdot \cdot \cdot .$$

EXAMPLE 1 Forming a Power Series

Use the function $f(x) = \sin x$ to form the Maclaurin series

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!}x^n = f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \frac{f^{(3)}(0)}{3!}x^3 + \frac{f^{(4)}(0)}{4!}x^4 + \cdot \cdot \cdot$$

and determine the interval of convergence.

Solution Successive differentiation of $f(x)$ yields

$$\begin{aligned}
f(x) &= \sin x & f(0) &= \sin 0 = 0\\
f'(x) &= \cos x & f'(0) &= \cos 0 = 1\\
f''(x) &= -\sin x & f''(0) &= -\sin 0 = 0\\
f^{(3)}(x) &= -\cos x & f^{(3)}(0) &= -\cos 0 = -1\\
f^{(4)}(x) &= \sin x & f^{(4)}(0) &= \sin 0 = 0\\
f^{(5)}(x) &= \cos x & f^{(5)}(0) &= \cos 0 = 1
\end{aligned}$$

and so on. The pattern repeats after the third derivative. So, the power series is as follows.

$$\begin{aligned}
\sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!}x^n &= f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \frac{f^{(3)}(0)}{3!}x^3 + \frac{f^{(4)}(0)}{4!}x^4 + \cdot \cdot \cdot\\
\sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)!} &= 0 + (1)x + \frac{0}{2!}x^2 + \frac{(-1)}{3!}x^3 + \frac{0}{4!}x^4 + \frac{1}{5!}x^5 + \frac{0}{6!}x^6\\
&\quad + \frac{(-1)}{7!}x^7 + \cdot \cdot \cdot\\
&= x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdot \cdot \cdot
\end{aligned}$$

By the Ratio Test, you can conclude that this series converges for all x.

Try It

Exploration A

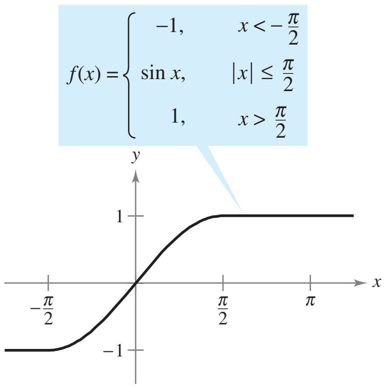


Figure 9.23

Notice that in Example 1 you cannot conclude that the power series converges to $\sin x$ for all x. You can simply conclude that the power series converges to some function, but you are not sure what function it is. This is a subtle, but important, point in dealing with Taylor or Maclaurin series. To persuade yourself that the series

$$f(c) + f'(c)(x - c) + \frac{f''(c)}{2!}(x - c)^2 + \cdot \cdot \cdot + \frac{f^{(n)}(c)}{n!}(x - c)^n + \cdot \cdot \cdot$$

might converge to a function other than f, remember that the derivatives are being evaluated at a single point. It can easily happen that another function will agree with the values of $f^{(n)}(x)$ when $x = c$ and disagree at other x-values. For instance, if you formed the power series (centered at 0) for the function shown in Figure 9.23, you would obtain the same series as in Example 1. You know that the series converges for all x, and yet it obviously cannot converge to both $f(x)$ and $\sin x$ for all x.

Let f have derivatives of all orders in an open interval I centered at c. The Taylor series for f may fail to converge for some x in I. Or, even if it is convergent, it may fail to have $f(x)$ as its sum. Nevertheless, Theorem 9.19 tells us that for each n,

$$f(x) = f(c) + f'(c)(x - c) + \frac{f''(c)}{2!}(x - c)^2 + \cdot \cdot \cdot + \frac{f^{(n)}(c)}{n!}(x - c)^n + R_n(x),$$

where

$$R_n(x) = \frac{f^{(n+1)}(z)}{(n + 1)!}(x - c)^{n+1}.$$

Note that in this remainder formula the particular value of z that makes the remainder formula true depends on the values of x and n. If $R_n \to 0$, then the following theorem tells us that the Taylor series for f actually converges to $f(x)$ for all x in I.

THEOREM 9.23 Convergence of Taylor Series

If $\lim\limits_{n\to\infty} R_n = 0$ for all x in the interval I, then the Taylor series for f converges and equals $f(x)$,

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(c)}{n!}(x - c)^n.$$

Proof For a Taylor series, the nth partial sum coincides with the nth Taylor polynomial. That is, $S_n(x) = P_n(x)$. Moreover, because

$$P_n(x) = f(x) - R_n(x)$$

it follows that

$$\begin{aligned} \lim_{n\to\infty} S_n(x) &= \lim_{n\to\infty} P_n(x) \\ &= \lim_{n\to\infty} [f(x) - R_n(x)] \\ &= f(x) - \lim_{n\to\infty} R_n(x). \end{aligned}$$

So, for a given x, the Taylor series (the sequence of partial sums) converges to $f(x)$ if and only if $R_n(x) \to 0$ as $n \to \infty$.

NOTE Stated another way, Theorem 9.23 says that a power series formed with Taylor coefficients $a_n = f^{(n)}(c)/n!$ converges to the function from which it was derived at precisely those values for which the remainder approaches 0 as $n \to \infty$.

In Example 1, you derived the power series from the sine function and you also concluded that the series converges to some function on the entire real line. In Example 2, you will see that the series actually converges to $\sin x$. The key observation is that although the value of z is not known, it is possible to obtain an upper bound for $|f^{(n+1)}(z)|$.

EXAMPLE 2 A Convergent Maclaurin Series

Show that the Maclaurin series for $f(x) = \sin x$ converges to $\sin x$ for all x.

Solution Using the result in Example 1, you need to show that

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdot \cdot \cdot + \frac{(-1)^n x^{2n+1}}{(2n+1)!} + \cdot \cdot \cdot$$

is true for all x. Because

$$f^{(n+1)}(x) = \pm \sin x$$

or

$$f^{(n+1)}(x) = \pm \cos x$$

you know that $|f^{(n+1)}(z)| \leq 1$ for every real number z. Therefore, for any fixed x, you can apply Taylor's Theorem (Theorem 9.19) to conclude that

$$0 \leq |R_n(x)| = \left| \frac{f^{(n+1)}(z)}{(n+1)!} x^{n+1} \right| \leq \frac{|x|^{n+1}}{(n+1)!}.$$

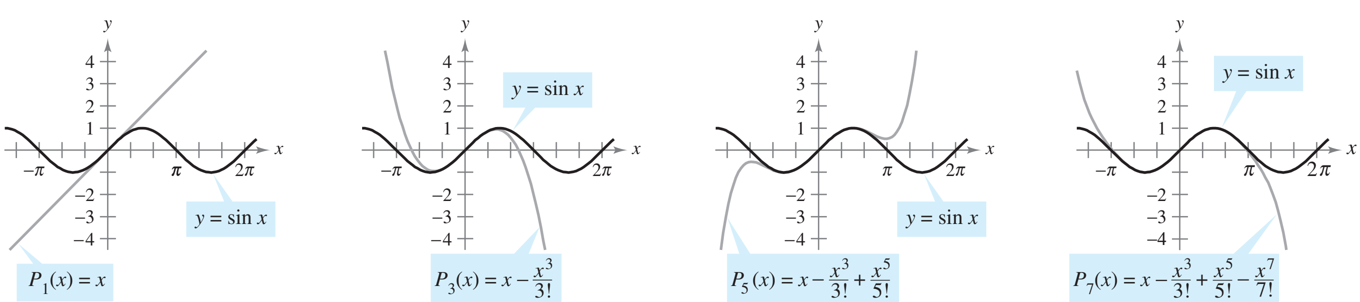
From the discussion in Section 9.1 regarding the relative rates of convergence of exponential and factorial sequences, it follows that for a fixed x

$$\lim_{n \to \infty} \frac{|x|^{n+1}}{(n+1)!} = 0.$$

Finally, by the Squeeze Theorem, it follows that for all x, $R_n(x) \to 0$ as $n \to \infty$. So, by Theorem 9.23, the Maclaurin series for $\sin x$ converges to $\sin x$ for all x.

Try It Exploration A Technology

Figure 9.24 visually illustrates the convergence of the Maclaurin series for $\sin x$ by comparing the graphs of the Maclaurin polynomials $P_1(x)$, $P_3(x)$, $P_5(x)$, and $P_7(x)$ with the graph of the sine function. Notice that as the degree of the polynomial increases, its graph more closely resembles that of the sine function.



As n increases, the graph of P_n more closely resembles the sine function.
Figure 9.24

The guidelines for finding a Taylor series for $f(x)$ at c are summarized below.

Guidelines for Finding a Taylor Series

1. Differentiate $f(x)$ several times and evaluate each derivative at c.

$$f(c), f'(c), f''(c), f'''(c), \cdot \cdot \cdot, f^{(n)}(c), \cdot \cdot \cdot$$

Try to recognize a pattern in these numbers.

2. Use the sequence developed in the first step to form the Taylor coefficients $a_n = f^{(n)}(c)/n!$, and determine the interval of convergence for the resulting power series

$$f(c) + f'(c)(x-c) + \frac{f''(c)}{2!}(x-c)^2 + \cdot \cdot \cdot + \frac{f^{(n)}(c)}{n!}(x-c)^n + \cdot \cdot \cdot .$$

3. Within this interval of convergence, determine whether or not the series converges to $f(x)$.

The direct determination of Taylor or Maclaurin coefficients using successive differentiation can be difficult, and the next example illustrates a shortcut for finding the coefficients indirectly—using the coefficients of a known Taylor or Maclaurin series.

EXAMPLE 3 Maclaurin Series for a Composite Function

Find the Maclaurin series for $f(x) = \sin x^2$.

Solution To find the coefficients for this Maclaurin series directly, you must calculate successive derivatives of $f(x) = \sin x^2$. By calculating just the first two,

$$f'(x) = 2x \cos x^2 \quad \text{and} \quad f''(x) = -4x^2 \sin x^2 + 2 \cos x^2$$

you can see that this task would be quite cumbersome. Fortunately, there is an alternative. First consider the Maclaurin series for $\sin x$ found in Example 1.

$$\begin{aligned} g(x) &= \sin x \\ &= x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdot \cdot \cdot \end{aligned}$$

Now, because $\sin x^2 = g(x^2)$, you can substitute x^2 for x in the series for $\sin x$ to obtain

$$\begin{aligned} \sin x^2 &= g(x^2) \\ &= x^2 - \frac{x^6}{3!} + \frac{x^{10}}{5!} - \frac{x^{14}}{7!} + \cdot \cdot \cdot . \end{aligned}$$

Try It   Exploration A

Be sure to understand the point illustrated in Example 3. Because direct computation of Taylor or Maclaurin coefficients can be tedious, the most practical way to find a Taylor or Maclaurin series is to develop power series for a *basic list* of elementary functions. From this list, you can determine power series for other functions by the operations of addition, subtraction, multiplication, division, differentiation, integration, or composition with known power series.

Binomial Series

Before presenting the basic list for elementary functions, you wll develop one more series—for a function of the form $f(x) = (1 + x)^k$. This produces the **binomial series.**

EXAMPLE 4 Binomial Series

Find the Maclaurin series for $f(x) = (1 + x)^k$ and determine its radius of convergence. Assume that k is not a positive integer.

Solution By successive differentiation, you have

$$\begin{aligned}
f(x) &= (1+x)^k & f(0) &= 1\\
f'(x) &= k(1+x)^{k-1} & f'(0) &= k\\
f''(x) &= k(k-1)(1+x)^{k-2} & f''(0) &= k(k-1)\\
f'''(x) &= k(k-1)(k-2)(1+x)^{k-3} & f'''(0) &= k(k-1)(k-2)\\
&\vdots & &\vdots\\
f^{(n)}(x) &= k \cdot \cdot \cdot (k-n+1)(1+x)^{k-n} & f^{(n)}(0) &= k(k-1) \cdot \cdot \cdot (k-n+1)
\end{aligned}$$

which produces the series

$$1 + kx + \frac{k(k-1)x^2}{2} + \cdot \cdot \cdot + \frac{k(k-1) \cdot \cdot \cdot (k-n+1)x^n}{n!} + \cdot \cdot \cdot .$$

Because $a_{n+1}/a_n \to 1$, you can apply the Ratio Test to conclude that the radius of convergence is $R = 1$. So, the series converges to some function in the interval $(-1, 1)$.

Exploration A

Note that Example 4 shows that the Taylor series for $(1 + x)^k$ converges to some function in the interval $(-1, 1)$. However, the example does not show that the series actually converges to $(1 + x)^k$. To do this, you could show that the remainder $R_n(x)$ converges to 0, as illustrated in Example 2.

EXAMPLE 5 Finding a Binomial Series

Find the power series for $f(x) = \sqrt[3]{1 + x}$.

Solution Using the binomial series

$$(1+x)^k = 1 + kx + \frac{k(k-1)x^2}{2!} + \frac{k(k-1)(k-2)x^3}{3!} + \cdot \cdot \cdot$$

let $k = \frac{1}{3}$ and write

$$(1+x)^{1/3} = 1 + \frac{x}{3} - \frac{2x^2}{3^2 2!} + \frac{2 \cdot 5x^3}{3^3 3!} - \frac{2 \cdot 5 \cdot 8x^4}{3^4 4!} + \cdot \cdot \cdot$$

which converges for $-1 \le x \le 1$.

Try It **Exploration A**

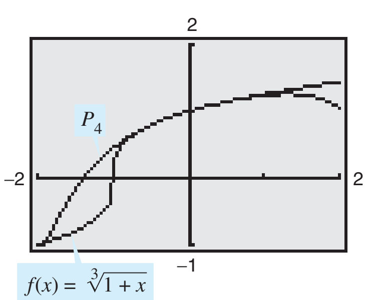


Figure 9.25

TECHNOLOGY Use a graphing utility to confirm the result in Example 5. When you graph the functions

$$f(x) = (1+x)^{1/3} \quad \text{and} \quad P_4(x) = 1 + \frac{x}{3} - \frac{x^2}{9} + \frac{5x^3}{81} - \frac{10x^4}{243}$$

in the same viewing window, you should obtain the result shown in Figure 9.25.

Deriving Taylor Series from a Basic List

The following list provides the power series for several elementary functions with the corresponding intervals of convergence.

Power Series for Elementary Functions

Function	*Interval of Convergence*
$\frac{1}{x} = 1 - (x-1) + (x-1)^2 - (x-1)^3 + (x-1)^4 - \cdots + (-1)^n(x-1)^n + \cdots$	$0 < x < 2$
$\frac{1}{1+x} = 1 - x + x^2 - x^3 + x^4 - x^5 + \cdots + (-1)^n x^n + \cdots$	$-1 < x < 1$
$\ln x = (x-1) - \frac{(x-1)^2}{2} + \frac{(x-1)^3}{3} - \frac{(x-1)^4}{4} + \cdots + \frac{(-1)^{n-1}(x-1)^n}{n} + \cdots$	$0 < x \le 2$
$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^5}{5!} + \cdots + \frac{x^n}{n!} + \cdots$	$-\infty < x < \infty$
$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \frac{x^9}{9!} - \cdots + \frac{(-1)^n x^{2n+1}}{(2n+1)!} + \cdots$	$-\infty < x < \infty$
$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \frac{x^8}{8!} - \cdots + \frac{(-1)^n x^{2n}}{(2n)!} + \cdots$	$-\infty < x < \infty$
$\arctan x = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \frac{x^9}{9} - \cdots + \frac{(-1)^n x^{2n+1}}{2n+1} + \cdots$	$-1 \le x \le 1$
$\arcsin x = x + \frac{x^3}{2 \cdot 3} + \frac{1 \cdot 3x^5}{2 \cdot 4 \cdot 5} + \frac{1 \cdot 3 \cdot 5x^7}{2 \cdot 4 \cdot 6 \cdot 7} + \cdots + \frac{(2n)!x^{2n+1}}{(2^n n!)^2(2n+1)} + \cdots$	$-1 \le x \le 1$
$(1+x)^k = 1 + kx + \frac{k(k-1)x^2}{2!} + \frac{k(k-1)(k-2)x^3}{3!} + \frac{k(k-1)(k-2)(k-3)x^4}{4!} + \cdots$	$-1 < x < 1$*

* *The convergence at $x = \pm 1$ depends on the value of k.*

NOTE The binomial series is valid for noninteger values of k. Moreover, if k happens to be a positive integer, the binomial series reduces to a simple binomial expansion.

EXAMPLE 6 Deriving a Power Series from a Basic List

Find the power series for $f(x) = \cos\sqrt{x}$.

Solution Using the power series

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \frac{x^8}{8!} - \cdots$$

you can replace x by $\sqrt{x}$ to obtain the series

$$\cos\sqrt{x} = 1 - \frac{x}{2!} + \frac{x^2}{4!} - \frac{x^3}{6!} + \frac{x^4}{8!} - \cdots.$$

This series converges for all x in the domain of $\cos\sqrt{x}$—that is, for $x \ge 0$.

Try It Exploration A

Power series can be multiplied and divided like polynomials. After finding the first few terms of the product (or quotient), you may be able to recognize a pattern.

EXAMPLE 7 Multiplication and Division of Power Series

Find the first three nonzero terms in each of the Maclaurin series.

a. $e^x \arctan x$ **b.** $\tan x$

Solution

a. Using the Maclaurin series for e^x and $\arctan x$ in the table, you have

$$e^x \arctan x = \left(1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots\right)\left(x - \frac{x^3}{3} + \frac{x^5}{5} - \cdots\right).$$

Multiply these expressions and collect like terms as you would for multiplying polynomials.

$$\begin{array}{rrrrrl}
1 + x & + \frac{1}{2}x^2 & + \frac{1}{6}x^3 & + \frac{1}{24}x^4 & & + \cdots \\
x & & - \frac{1}{3}x^3 & & + \frac{1}{5}x^5 & - \cdots \\
\hline
x + & x^2 & + \frac{1}{2}x^3 & + \frac{1}{6}x^4 & + \frac{1}{24}x^5 & + \cdots \\
& & - \frac{1}{3}x^3 & - \frac{1}{3}x^4 & - \frac{1}{6}x^5 & - \cdots \\
& & & & + \frac{1}{5}x^5 & + \cdots \\
\hline
x + & x^2 & + \frac{1}{6}x^3 & - \frac{1}{6}x^4 & + \frac{3}{40}x^5 & + \cdots
\end{array}$$

So, $e^x \arctan x = x + x^2 + \frac{1}{6}x^3 + \cdots$.

b. Using the Maclaurin series for $\sin x$ and $\cos x$ in the table, you have

$$\tan x = \frac{\sin x}{\cos x} = \frac{x - \dfrac{x^3}{3!} + \dfrac{x^5}{5!} - \cdots}{1 - \dfrac{x^2}{2!} + \dfrac{x^4}{4!} - \cdots}.$$

Divide using long division.

$$\begin{array}{r|rrl}
& x + \frac{1}{3}x^3 & + \frac{2}{15}x^5 & + \cdots \\
\hline
1 - \frac{1}{2}x^2 + \frac{1}{24}x^4 - \cdots & x - \frac{1}{6}x^3 & + \frac{1}{120}x^5 & - \cdots \\
& x - \frac{1}{2}x^3 & + \frac{1}{24}x^5 & - \cdots \\
\cline{2-4}
& \frac{1}{3}x^3 & - \frac{1}{30}x^5 & + \cdots \\
& \frac{1}{3}x^3 & - \frac{1}{6}x^5 & + \cdots \\
\cline{2-4}
& & \frac{2}{15}x^5 & + \cdots
\end{array}$$

So, $\tan x = x + \frac{1}{3}x^3 + \frac{2}{15}x^5 + \cdots$.

Try It Exploration A

EXAMPLE 8 A Power Series for $\sin^2 x$

Find the power series for $f(x) = \sin^2 x$.

Solution Consider rewriting $\sin^2 x$ as follows.

$$\sin^2 x = \frac{1 - \cos 2x}{2} = \frac{1}{2} - \frac{\cos 2x}{2}$$

Now, use the series for $\cos x$.

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \frac{x^8}{8!} - \cdots$$

$$\cos 2x = 1 - \frac{2^2}{2!}x^2 + \frac{2^4}{4!}x^4 - \frac{2^6}{6!}x^6 + \frac{2^8}{8!}x^8 - \cdots$$

$$-\frac{1}{2}\cos 2x = -\frac{1}{2} + \frac{2}{2!}x^2 - \frac{2^3}{4!}x^4 + \frac{2^5}{6!}x^6 - \frac{2^7}{8!}x^8 + \cdots$$

$$\sin^2 x = \frac{1}{2} - \frac{1}{2}\cos 2x = \frac{1}{2} - \frac{1}{2} + \frac{2}{2!}x^2 - \frac{2^3}{4!}x^4 + \frac{2^5}{6!}x^6 - \frac{2^7}{8!}x^8 + \cdots$$

$$= \frac{2}{2!}x^2 - \frac{2^3}{4!}x^4 + \frac{2^5}{6!}x^6 - \frac{2^7}{8!}x^8 + \cdots$$

This series converges for $-\infty < x < \infty$.

Try It   Exploration A

As mentioned in the preceding section, power series can be used to obtain tables of values of transcendental functions. They are also useful for estimating the values of definite integrals for which antiderivatives cannot be found. The next example demonstrates this use.

EXAMPLE 9 Power Series Approximation of a Definite Integral

Use a power series to approximate

$$\int_0^1 e^{-x^2}\,dx$$

with an error of less than 0.01.

Solution Replacing x with $-x^2$ in the series for e^x produces the following.

$$e^{-x^2} = 1 - x^2 + \frac{x^4}{2!} - \frac{x^6}{3!} + \frac{x^8}{4!} - \cdots$$

$$\int_0^1 e^{-x^2}\,dx = \left[x - \frac{x^3}{3} + \frac{x^5}{5 \cdot 2!} - \frac{x^7}{7 \cdot 3!} + \frac{x^9}{9 \cdot 4!} - \cdots\right]_0^1$$

$$= 1 - \frac{1}{3} + \frac{1}{10} - \frac{1}{42} + \frac{1}{216} - \cdots$$

Summing the first four terms, you have

$$\int_0^1 e^{-x^2}\,dx \approx 0.74$$

which, by the Alternating Series Test, has an error of less than $\frac{1}{216} \approx 0.005$.

Try It   Exploration A   Open Exploration

Section 10.1 Conics and Calculus

- **Understand the definition of a conic section.**
- **Analyze and write equations of parabolas using properties of parabolas.**
- **Analyze and write equations of ellipses using properties of ellipses.**
- **Analyze and write equations of hyperbolas using properties of hyperbolas.**

HYPATIA (370–415 A.D.)

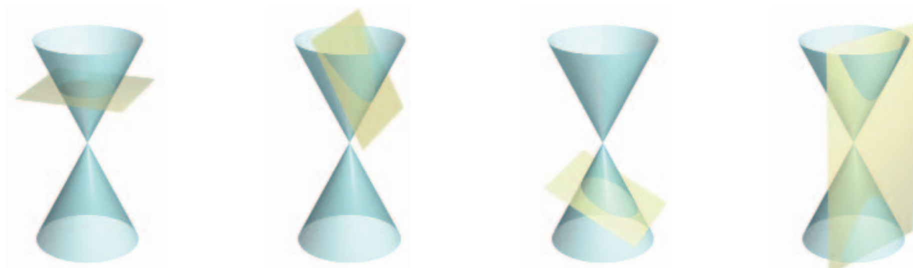
The Greeks discovered conic sections sometime between 600 and 300 B.C. By the beginning of the Alexandrian period, enough was known about conics for Apollonius (262–190 B.C.) to produce an eight-volume work on the subject. Later, toward the end of the Alexandrian period, Hypatia wrote a textbook entitled *On the Conics of Apollonius*. Her death marked the end of major mathematical discoveries in Europe for several hundred years.

The early Greeks were largely concerned with the geometric properties of conics. It was not until 1900 years later, in the early seventeenth century, that the broader applicability of conics became apparent. Conics then played a prominent role in the development of calculus.

MathBio

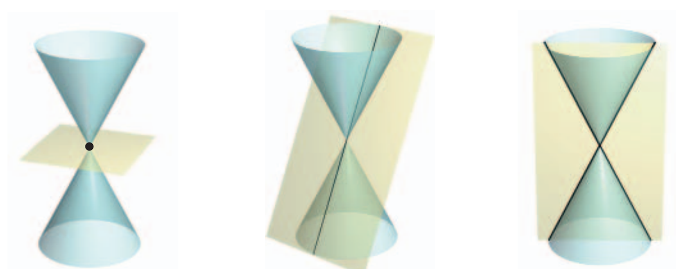
Conic Sections

Each **conic section** (or simply **conic**) can be described as the intersection of a plane and a double-napped cone. Notice in Figure 10.1 that for the four basic conics, the intersecting plane does not pass through the vertex of the cone. When the plane passes through the vertex, the resulting figure is a **degenerate conic,** as shown in Figure 10.2.



Circle Parabola Ellipse Hyperbola
Conic sections
Figure 10.1

Animation



Point Line Two intersecting lines
Degenerate conics
Figure 10.2

Animation

There are several ways to study conics. You could begin as the Greeks did by defining the conics in terms of the intersections of planes and cones, or you could define them algebraically in terms of the general second-degree equation

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0. \qquad \text{General second-degree equation}$$

However, a third approach, in which each of the conics is defined as a **locus** (collection) of points satisfying a certain geometric property, works best. For example, a circle can be defined as the collection of all points (x, y) that are equidistant from a fixed point (h, k). This locus definition easily produces the standard equation of a circle

$$(x - h)^2 + (y - k)^2 = r^2. \qquad \text{Standard equation of a circle}$$

FOR FURTHER INFORMATION To learn more about the mathematical activities of Hypatia, see the article "Hypatia and Her Mathematics" by Michael A. B. Deakin in *The American Mathematical Monthly*.

MathArticle

Parabolas

A **parabola** is the set of all points (x, y) that are equidistant from a fixed line called the **directrix** and a fixed point called the **focus** not on the line. The midpoint between the focus and the directrix is the **vertex,** and the line passing through the focus and the vertex is the **axis** of the parabola. Note in Figure 10.3 that a parabola is symmetric with respect to its axis.

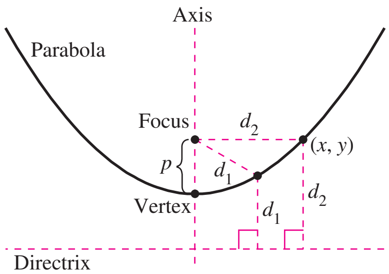


Figure 10.3

THEOREM 10.1 Standard Equation of a Parabola

The **standard form** of the equation of a parabola with vertex (h, k) and directrix $y = k - p$ is

$$(x - h)^2 = 4p(y - k). \qquad \text{Vertical axis}$$

For directrix $x = h - p$, the equation is

$$(y - k)^2 = 4p(x - h). \qquad \text{Horizontal axis}$$

The focus lies on the axis p units (*directed distance*) from the vertex. The coordinates of the focus are as follows.

$$(h, k + p) \qquad \text{Vertical axis}$$
$$(h + p, k) \qquad \text{Horizontal axis}$$

EXAMPLE 1 Finding the Focus of a Parabola

Find the focus of the parabola given by $y = -\frac{1}{2}x^2 - x + \frac{1}{2}$.

Solution To find the focus, convert to standard form by completing the square.

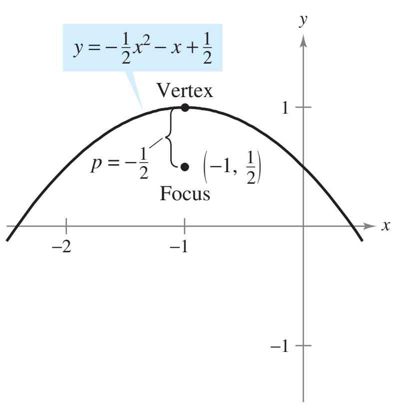
$$\begin{aligned}
y &= \tfrac{1}{2} - x - \tfrac{1}{2}x^2 && \text{Rewrite original equation.}\\
y &= \tfrac{1}{2}(1 - 2x - x^2) && \text{Factor out } \tfrac{1}{2}.\\
2y &= 1 - 2x - x^2 && \text{Multiply each side by 2.}\\
2y &= 1 - (x^2 + 2x) && \text{Group terms.}\\
2y &= 2 - (x^2 + 2x + 1) && \text{Add and subtract 1 on right side.}\\
x^2 + 2x + 1 &= -2y + 2 &&\\
(x + 1)^2 &= -2(y - 1) && \text{Write in standard form.}
\end{aligned}$$

Comparing this equation with $(x - h)^2 = 4p(y - k)$, you can conclude that

$$h = -1, \qquad k = 1, \qquad \text{and} \qquad p = -\tfrac{1}{2}.$$

Because p is negative, the parabola opens downward, as shown in Figure 10.4. So, the focus of the parabola is p units from the vertex, or

$$(h, k + p) = \left(-1, \tfrac{1}{2}\right). \qquad \text{Focus}$$



Parabola with a vertical axis, $p < 0$
Figure 10.4

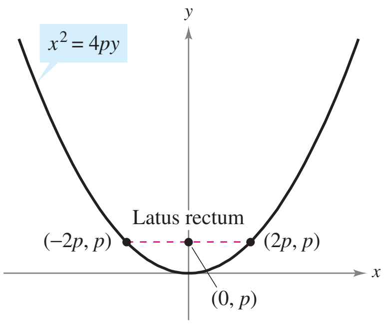
Editable Graph

Try It

Exploration A

Exploration B

A line segment that passes through the focus of a parabola and has endpoints on the parabola is called a **focal chord.** The specific focal chord perpendicular to the axis of the parabola is the **latus rectum.** The next example shows how to determine the length of the latus rectum and the length of the corresponding intercepted arc.



Length of latus rectum: $4p$
Figure 10.5

EXAMPLE 2 Focal Chord Length and Arc Length

Find the length of the latus rectum of the parabola given by $x^2 = 4py$. Then find the length of the parabolic arc intercepted by the latus rectum.

Solution Because the latus rectum passes through the focus $(0, p)$ and is perpendicular to the y-axis, the coordinates of its endpoints are $(-x, p)$ and (x, p). Substituting p for y in the equation of the parabola produces

$$x^2 = 4p(p) \quad \Longrightarrow \quad x = \pm 2p.$$

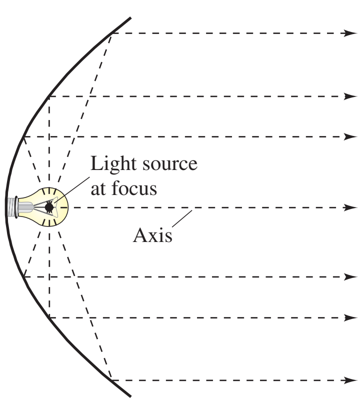
So, the endpoints of the latus rectum are $(-2p, p)$ and $(2p, p)$, and you can conclude that its length is $4p$, as shown in Figure 10.5. In contrast, the length of the intercepted arc is

$$\begin{aligned}
s &= \int_{-2p}^{2p} \sqrt{1 + (y')^2}\, dx && \text{Use arc length formula.}\\
&= 2\int_0^{2p} \sqrt{1 + \left(\frac{x}{2p}\right)^2}\, dx && y = \frac{x^2}{4p} \Longrightarrow y' = \frac{x}{2p}\\
&= \frac{1}{p}\int_0^{2p} \sqrt{4p^2 + x^2}\, dx && \text{Simplify.}\\
&= \frac{1}{2p}\left[x\sqrt{4p^2 + x^2} + 4p^2 \ln\left|x + \sqrt{4p^2 + x^2}\right|\right]_0^{2p} && \text{Theorem 8.2}\\
&= \frac{1}{2p}\left[2p\sqrt{8p^2} + 4p^2 \ln\left(2p + \sqrt{8p^2}\right) - 4p^2 \ln(2p)\right]\\
&= 2p\left[\sqrt{2} + \ln\left(1 + \sqrt{2}\right)\right]\\
&\approx 4.59p.
\end{aligned}$$

Exploration A Open Exploration

One widely used property of a parabola is its reflective property. In physics, a surface is called **reflective** if the tangent line at any point on the surface makes equal angles with an incoming ray and the resulting outgoing ray. The angle corresponding to the incoming ray is the **angle of incidence,** and the angle corresponding to the outgoing ray is the **angle of reflection.** One example of a reflective surface is a flat mirror.

Another type of reflective surface is that formed by revolving a parabola about its axis. A special property of parabolic reflectors is that they allow us to direct all incoming rays parallel to the axis through the focus of the parabola—this is the principle behind the design of the parabolic mirrors used in reflecting telescopes. Conversely, all light rays emanating from the focus of a parabolic reflector used in a flashlight are parallel, as shown in Figure 10.6.



Parabolic reflector: light is reflected in parallel rays.
Figure 10.6

THEOREM 10.2 Reflective Property of a Parabola

Let P be a point on a parabola. The tangent line to the parabola at the point P makes equal angles with the following two lines.

1. The line passing through P and the focus
2. The line passing through P parallel to the axis of the parabola

Nicolaus Copernicus (1473–1543)

Copernicus began to study planetary motion when asked to revise the calendar. At that time, the exact length of the year could not be accurately predicted using the theory that Earth was the center of the universe.

MathBio

Ellipses

More than a thousand years after the close of the Alexandrian period of Greek mathematics, Western civilization finally began a Renaissance of mathematical and scientific discovery. One of the principal figures in this rebirth was the Polish astronomer Nicolaus Copernicus. In his work *On the Revolutions of the Heavenly Spheres*, Copernicus claimed that all of the planets, including Earth, revolved about the sun in circular orbits. Although some of Copernicus's claims were invalid, the controversy set off by his heliocentric theory motivated astronomers to search for a mathematical model to explain the observed movements of the sun and planets. The first to find an accurate model was the German astronomer Johannes Kepler (1571–1630). Kepler discovered that the planets move about the sun in elliptical orbits, with the sun not as the center but as a focal point of the orbit.

The use of ellipses to explain the movements of the planets is only one of many practical and aesthetic uses. As with parabolas, you will begin your study of this second type of conic by defining it as a locus of points. Now, however, *two* focal points are used rather than one.

An **ellipse** is the set of all points (x, y) the sum of whose distances from two distinct fixed points called **foci** is constant. (See Figure 10.7.) The line through the foci intersects the ellipse at two points, called the **vertices.** The chord joining the vertices is the **major axis,** and its midpoint is the **center** of the ellipse. The chord perpendicular to the major axis at the center is the **minor axis** of the ellipse. (See Figure 10.8.)

FOR FURTHER INFORMATION To learn about how an ellipse may be "exploded" into a parabola, see the article "Exploding the Ellipse" by Arnold Good in *Mathematics Teacher*.

MathArticle

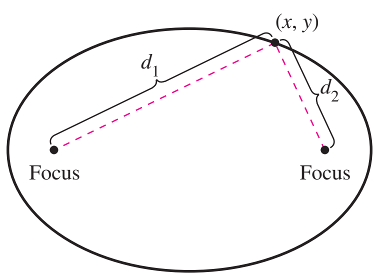


Figure 10.7

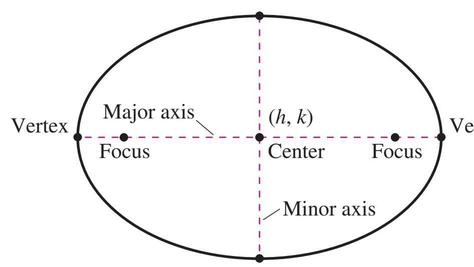


Figure 10.8

THEOREM 10.3 Standard Equation of an Ellipse

The standard form of the equation of an ellipse with center (h, k) and major and minor axes of lengths $2a$ and $2b$, where $a > b$, is

$$\frac{(x-h)^2}{a^2} + \frac{(y-k)^2}{b^2} = 1 \qquad \text{Major axis is horizontal.}$$

or

$$\frac{(x-h)^2}{b^2} + \frac{(y-k)^2}{a^2} = 1. \qquad \text{Major axis is vertical.}$$

The foci lie on the major axis, c units from the center, with $c^2 = a^2 - b^2$.

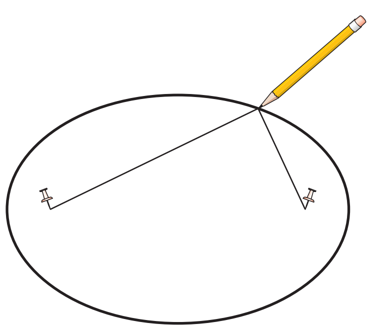


Figure 10.9

Animation

NOTE You can visualize the definition of an ellipse by imagining two thumbtacks placed at the foci, as shown in Figure 10.9. If the ends of a fixed length of string are fastened to the thumbtacks and the string is drawn taut with a pencil, the path traced by the pencil will be an ellipse.

EXAMPLE 3 Completing the Square

Find the center, vertices, and foci of the ellipse given by

$$4x^2 + y^2 - 8x + 4y - 8 = 0.$$

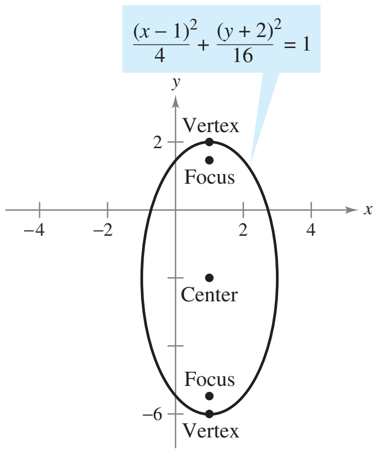
Solution By completing the square, you can write the original equation in standard form.

$$\begin{aligned}
4x^2 + y^2 - 8x + 4y - 8 &= 0 && \text{Write original equation.}\\
4x^2 - 8x + y^2 + 4y &= 8 &&\\
4(x^2 - 2x + 1) + (y^2 + 4y + 4) &= 8 + 4 + 4 &&\\
4(x - 1)^2 + (y + 2)^2 &= 16 &&\\
\frac{(x-1)^2}{4} + \frac{(y+2)^2}{16} &= 1 && \text{Write in standard form.}
\end{aligned}$$

So, the major axis is parallel to the y-axis, where $h = 1$, $k = -2$, $a = 4$, $b = 2$, and $c = \sqrt{16 - 4} = 2\sqrt{3}$. So, you obtain the following.

Center:	$(1, -2)$	(h, k)
Vertices:	$(1, -6)$ and $(1, 2)$	$(h, k \pm a)$
Foci:	$\left(1, -2 - 2\sqrt{3}\right)$ and $\left(1, -2 + 2\sqrt{3}\right)$	$(h, k \pm c)$

The graph of the ellipse is shown in Figure 10.10.



Ellipse with a vertical major axis
Figure 10.10

Editable Graph

Try It Exploration A Exploration B

NOTE If the constant term $F = -8$ in the equation in Example 3 had been greater than or equal to 8, you would have obtained one of the following degenerate cases.

1. $F = 8$, single point, $(1, -2)$: $\dfrac{(x-1)^2}{4} + \dfrac{(y+2)^2}{16} = 0$

2. $F > 8$, no solution points: $\dfrac{(x-1)^2}{4} + \dfrac{(y+2)^2}{16} < 0$

EXAMPLE 4 The Orbit of the Moon

The moon orbits Earth in an elliptical path with the center of Earth at one focus, as shown in Figure 10.11. The major and minor axes of the orbit have lengths of 768,800 kilometers and 767,640 kilometers, respectively. Find the greatest and least distances (the apogee and perigee) from Earth's center to the moon's center.

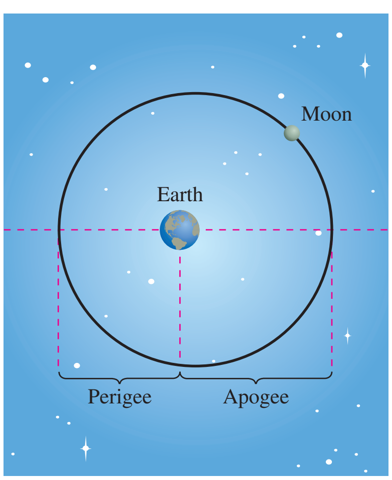


Figure 10.11

Solution Begin by solving for a and b.

$$\begin{aligned}
2a &= 768{,}800 && \text{Length of major axis}\\
a &= 384{,}400 && \text{Solve for } a.\\
2b &= 767{,}640 && \text{Length of minor axis}\\
b &= 383{,}820 && \text{Solve for } b.
\end{aligned}$$

Now, using these values, you can solve for c as follows.

$$c = \sqrt{a^2 - b^2} \approx 21{,}108$$

The greatest distance between the center of Earth and the center of the moon is $a + c \approx 405{,}508$ kilometers, and the least distance is $a - c \approx 363{,}292$ kilometers.

Try It Exploration A

FOR FURTHER INFORMATION For more information on some uses of the reflective properties of conics, see the article "The Geometry of Microwave Antennas" by William R. Parzynski in *Mathematics Teacher*.

MathArticle

Theorem 10.2 presented a reflective property of parabolas. Ellipses have a similar reflective property. You are asked to prove the following theorem in Exercise 110.

THEOREM 10.4 Reflective Property of an Ellipse

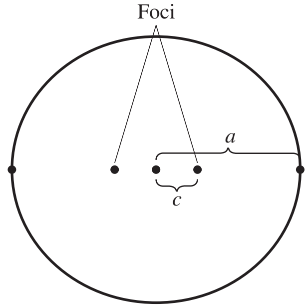
Let P be a point on an ellipse. The tangent line to the ellipse at point P makes equal angles with the lines through P and the foci.

One of the reasons that astronomers had difficulty in detecting that the orbits of the planets are ellipses is that the foci of the planetary orbits are relatively close to the center of the sun, making the orbits nearly circular. To measure the ovalness of an ellipse, you can use the concept of **eccentricity.**

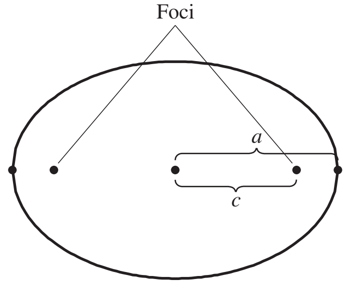
Definition of Eccentricity of an Ellipse

The **eccentricity** e of an ellipse is given by the ratio

$$e = \frac{c}{a}.$$



(a) $\frac{c}{a}$ is small.



(b) $\frac{c}{a}$ is close to 1.

Eccentricity is the ratio $\frac{c}{a}$.
Figure 10.12

To see how this ratio is used to describe the shape of an ellipse, note that because the foci of an ellipse are located along the major axis between the vertices and the center, it follows that

$$0 < c < a.$$

For an ellipse that is nearly circular, the foci are close to the center and the ratio c/a is small, and for an elongated ellipse, the foci are close to the vertices and the ratio is close to 1, as shown in Figure 10.12. Note that $0 < e < 1$ for every ellipse.

The orbit of the moon has an eccentricity of $e = 0.0549$, and the eccentricities of the nine planetary orbits are as follows.

Mercury:	$e = 0.2056$	Saturn:	$e = 0.0542$
Venus:	$e = 0.0068$	Uranus:	$e = 0.0472$
Earth:	$e = 0.0167$	Neptune:	$e = 0.0086$
Mars:	$e = 0.0934$	Pluto:	$e = 0.2488$
Jupiter:	$e = 0.0484$		

You can use integration to show that the area of an ellipse is $A = \pi ab$. For instance, the area of the ellipse

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

is given by

$$A = 4\int_0^a \frac{b}{a}\sqrt{a^2 - x^2}\,dx$$

$$= \frac{4b}{a}\int_0^{\pi/2} a^2 \cos^2\theta\,d\theta. \qquad \text{Trigonometric substitution } x = a\sin\theta.$$

However, it is not so simple to find the *circumference* of an ellipse. The next example shows how to use eccentricity to set up an "elliptic integral" for the circumference of an ellipse.

EXAMPLE 5 Finding the Circumference of an Ellipse

Show that the circumference of the ellipse $(x^2/a^2) + (y^2/b^2) = 1$ is

$$4a\int_0^{\pi/2} \sqrt{1 - e^2 \sin^2 \theta}\, d\theta. \qquad e = \frac{c}{a}$$

Solution Because the given ellipse is symmetric with respect to both the x-axis and the y-axis, you know that its circumference C is four times the arc length of $y = (b/a)\sqrt{a^2 - x^2}$ in the first quadrant. The function y is differentiable for all x in the interval $[0, a]$ except at $x = a$. So, the circumference is given by the improper integral

$$C = \lim_{d \to a} 4\int_0^d \sqrt{1 + (y')^2}\, dx = 4\int_0^a \sqrt{1 + (y')^2}\, dx = 4\int_0^a \sqrt{1 + \frac{b^2x^2}{a^2(a^2 - x^2)}}\, dx.$$

Using the trigonometric substitution $x = a \sin \theta$, you obtain

$$\begin{aligned} C &= 4\int_0^{\pi/2} \sqrt{1 + \frac{b^2 \sin^2 \theta}{a^2 \cos^2 \theta}}\,(a \cos \theta)\, d\theta \\ &= 4\int_0^{\pi/2} \sqrt{a^2 \cos^2 \theta + b^2 \sin^2 \theta}\, d\theta \\ &= 4\int_0^{\pi/2} \sqrt{a^2(1 - \sin^2 \theta) + b^2 \sin^2 \theta}\, d\theta \\ &= 4\int_0^{\pi/2} \sqrt{a^2 - (a^2 - b^2)\sin^2 \theta}\, d\theta. \end{aligned}$$

Because $e^2 = c^2/a^2 = (a^2 - b^2)/a^2$, you can rewrite this integral as

$$C = 4a\int_0^{\pi/2} \sqrt{1 - e^2 \sin^2 \theta}\, d\theta.$$

Area and Circumference of an Ellipse

In his work with elliptic orbits in the early 1600's, Johannes Kepler successfully developed a formula for the area of an ellipse, $A = \pi ab$. He was less successful in developing a formula for the circumference of an ellipse, however; the best he could do was to give the approximate formula $C = \pi(a + b)$.

Exploration A | Exploration B | Open Exploration

A great deal of time has been devoted to the study of elliptic integrals. Such integrals generally do not have elementary antiderivatives. To find the circumference of an ellipse, you must usually resort to an approximation technique.

EXAMPLE 6 Approximating the Value of an Elliptic Integral

Use the elliptic integral in Example 5 to approximate the circumference of the ellipse

$$\frac{x^2}{25} + \frac{y^2}{16} = 1.$$

Solution Because $e^2 = c^2/a^2 = (a^2 - b^2)/a^2 = 9/25$, you have

$$C = (4)(5)\int_0^{\pi/2} \sqrt{1 - \frac{9 \sin^2 \theta}{25}}\, d\theta.$$

Applying Simpson's Rule with $n = 4$ produces

$$\begin{aligned} C &\approx 20\left(\frac{\pi}{6}\right)\left(\frac{1}{4}\right)[1 + 4(0.9733) + 2(0.9055) + 4(0.8323) + 0.8] \\ &\approx 28.36. \end{aligned}$$

So, the ellipse has a circumference of about 28.36 units, as shown in Figure 10.13.

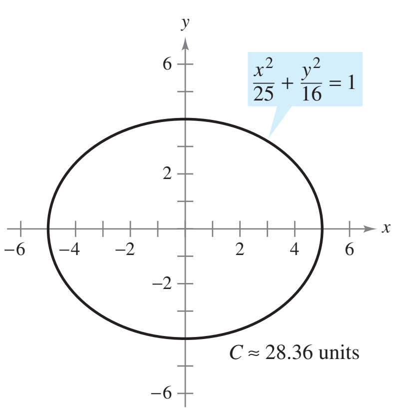


Figure 10.13

Try It | Exploration A

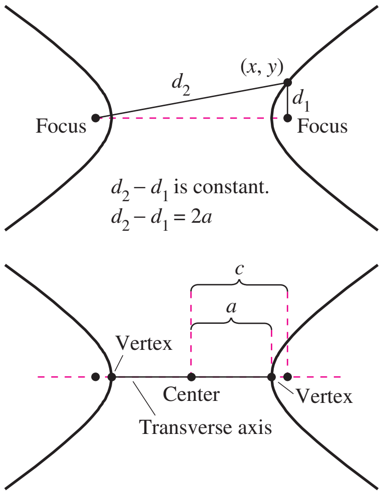


Figure 10.14

Hyperbolas

The definition of a hyperbola is similar to that of an ellipse. For an ellipse, the *sum* of the distances between the foci and a point on the ellipse is fixed, whereas for a hyperbola, the absolute value of the *difference* between these distances is fixed.

A **hyperbola** is the set of all points (x, y) for which the absolute value of the difference between the distances from two distinct fixed points called **foci** is constant. (See Figure 10.14.) The line through the two foci intersects a hyperbola at two points called the **vertices.** The line segment connecting the vertices is the **transverse axis,** and the midpoint of the transverse axis is the **center** of the hyperbola. One distinguishing feature of a hyperbola is that its graph has two separate *branches*.

THEOREM 10.5 Standard Equation of a Hyperbola

The standard form of the equation of a hyperbola with center at (h, k) is

$$\frac{(x-h)^2}{a^2} - \frac{(y-k)^2}{b^2} = 1 \qquad \text{Transverse axis is horizontal.}$$

or

$$\frac{(y-k)^2}{a^2} - \frac{(x-h)^2}{b^2} = 1. \qquad \text{Transverse axis is vertical.}$$

The vertices are a units from the center, and the foci are c units from the center, where, $c^2 = a^2 + b^2$.

NOTE The constants a, b, and c do not have the same relationship for hyperbolas as they do for ellipses. For hyperbolas, $c^2 = a^2 + b^2$, but for ellipses, $c^2 = a^2 - b^2$.

An important aid in sketching the graph of a hyperbola is the determination of its **asymptotes,** as shown in Figure 10.15. Each hyperbola has two asymptotes that intersect at the center of the hyperbola. The asymptotes pass through the vertices of a rectangle of dimensions $2a$ by $2b$, with its center at (h, k). The line segment of length $2b$ joining $(h, k + b)$ and $(h, k - b)$ is referred to as the **conjugate axis** of the hyperbola.

THEOREM 10.6 Asymptotes of a Hyperbola

For a *horizontal* transverse axis, the equations of the asymptotes are

$$y = k + \frac{b}{a}(x - h) \qquad \text{and} \qquad y = k - \frac{b}{a}(x - h).$$

For a *vertical* transverse axis, the equations of the asymptotes are

$$y = k + \frac{a}{b}(x - h) \qquad \text{and} \qquad y = k - \frac{a}{b}(x - h).$$

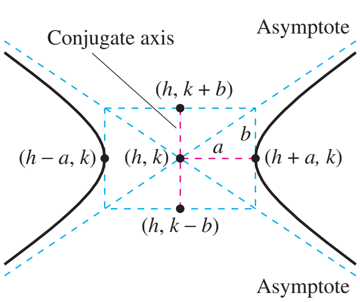


Figure 10.15

In Figure 10.15 you can see that the asymptotes coincide with the diagonals of the rectangle with dimensions $2a$ and $2b$, centered at (h, k). This provides you with a quick means of sketching the asymptotes, which in turn aids in sketching the hyperbola.

EXAMPLE 7 Using Asymptotes to Sketch a Hyperbola

Sketch the graph of the hyperbola whose equation is $4x^2 - y^2 = 16$.

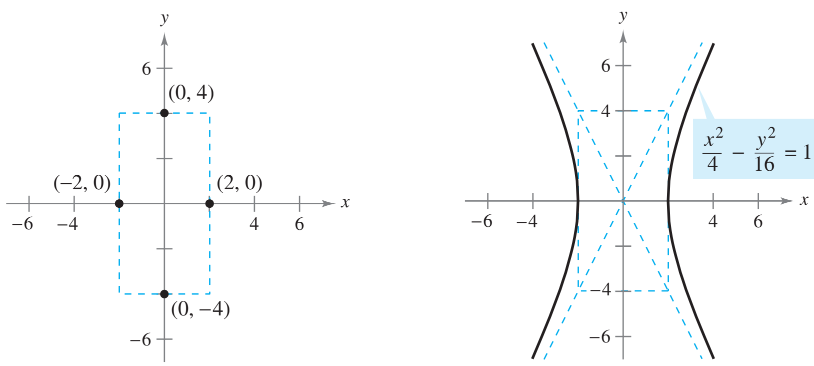
TECHNOLOGY You can use a graphing utility to verify the graph obtained in Example 7 by solving the original equation for y and graphing the following equations.

$$y_1 = \sqrt{4x^2 - 16}$$
$$y_2 = -\sqrt{4x^2 - 16}$$

Solution Begin by rewriting the equation in standard form.

$$\frac{x^2}{4} - \frac{y^2}{16} = 1$$

The transverse axis is horizontal and the vertices occur at $(-2, 0)$ and $(2, 0)$. The ends of the conjugate axis occur at $(0, -4)$ and $(0, 4)$. Using these four points, you can sketch the rectangle shown in Figure 10.16(a). By drawing the asymptotes through the corners of this rectangle, you can complete the sketch as shown in Figure 10.16(b).



(a) (b)

Figure 10.16

Editable Graph

Try It | Exploration A | Exploration B | Exploration C

Exploration D | Exploration E | Open Exploration

Definition of Eccentricity of a Hyperbola

The **eccentricity** e of a hyperbola is given by the ratio

$$e = \frac{c}{a}.$$

As with an ellipse, the **eccentricity** of a hyperbola is $e = c/a$. Because $c > a$ for hyperbolas, it follows that $e > 1$ for hyperbolas. If the eccentricity is large, the branches of the hyperbola are nearly flat. If the eccentricity is close to 1, the branches of the hyperbola are more pointed, as shown in Figure 10.17.

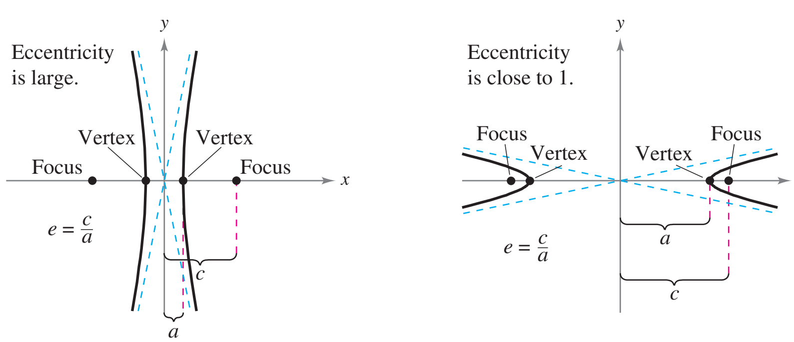
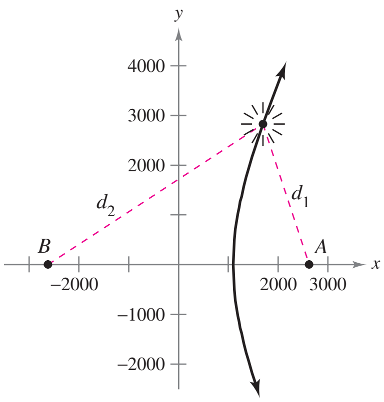


Figure 10.17

The following application was developed during World War II. It shows how the properties of hyperbolas can be used in radar and other detection systems.

EXAMPLE 8 A Hyperbolic Detection System

Two microphones, 1 mile apart, record an explosion. Microphone A receives the sound 2 seconds before microphone B. Where was the explosion?



$2c = 5280$
$d_2 - d_1 = 2a = 2200$
Figure 10.18

Solution Assuming that sound travels at 1100 feet per second, you know that the explosion took place 2200 feet farther from B than from A, as shown in Figure 10.18. The locus of all points that are 2200 feet closer to A than to B is one branch of the hyperbola $(x^2/a^2) - (y^2/b^2) = 1$, where

$$c = \frac{1 \text{ mile}}{2} = \frac{5280 \text{ ft}}{2} = 2640 \text{ feet}$$

and

$$a = \frac{2200 \text{ ft}}{2} = 1100 \text{ feet.}$$

Because $c^2 = a^2 + b^2$, it follows that

$$\begin{aligned} b^2 &= c^2 - a^2 \\ &= 5{,}759{,}600 \end{aligned}$$

and you can conclude that the explosion occurred somewhere on the right branch of the hyperbola given by

$$\frac{x^2}{1{,}210{,}000} - \frac{y^2}{5{,}759{,}600} = 1.$$

Try It Exploration A

In Example 8, you were able to determine only the hyperbola on which the explosion occurred, but not the exact location of the explosion. If, however, you had received the sound at a third position C, then two other hyperbolas would be determined. The exact location of the explosion would be the point at which these three hyperbolas intersect.

Another interesting application of conics involves the orbits of comets in our solar system. Of the 610 comets identified prior to 1970, 245 have elliptical orbits, 295 have parabolic orbits, and 70 have hyperbolic orbits. The center of the sun is a focus of each orbit, and each orbit has a vertex at the point at which the comet is closest to the sun. Undoubtedly, many comets with parabolic or hyperbolic orbits have not been identified—such comets pass through our solar system only once. Only comets with elliptical orbits such as Halley's comet remain in our solar system.

The type of orbit for a comet can be determined as follows.

1. Ellipse: $v < \sqrt{2GM/p}$
2. Parabola: $v = \sqrt{2GM/p}$
3. Hyperbola: $v > \sqrt{2GM/p}$

In these three formulas, p is the distance between one vertex and one focus of the comet's orbit (in meters), v is the velocity of the comet at the vertex (in meters per second), $M \approx 1.989 \times 10^{30}$ kilograms is the mass of the sun, and $G \approx 6.67 \times 10^{-8}$ cubic meters per kilogram-second squared is the gravitational constant. View the video for more information about a comet with an elliptical orbit.

Video

Caroline Herschel (1750–1848)

The first woman to be credited with detecting a new comet was the English astronomer Caroline Herschel. During her life, Caroline Herschel discovered a total of eight new comets.

MathBio

Section 10.2

Plane Curves and Parametric Equations

- **Sketch the graph of a curve given by a set of parametric equations.**
- **Eliminate the parameter in a set of parametric equations.**
- **Find a set of parametric equations to represent a curve.**
- **Understand two classic calculus problems, the tautochrone and brachistochrone problems.**

Plane Curves and Parametric Equations

Until now, you have been representing a graph by a single equation involving *two* variables. In this section you will study situations in which *three* variables are used to represent a curve in the plane.

Consider the path followed by an object that is propelled into the air at an angle of 45°. If the initial velocity of the object is 48 feet per second, the object travels the parabolic path given by

$$y = -\frac{x^2}{72} + x \qquad \text{Rectangular equation}$$

as shown in Figure 10.19. However, this equation does not tell the whole story. Although it does tell you *where* the object has been, it doesn't tell you *when* the object was at a given point (x, y). To determine this time, you can introduce a third variable t, called a **parameter.** By writing both x and y as functions of t, you obtain the **parametric equations**

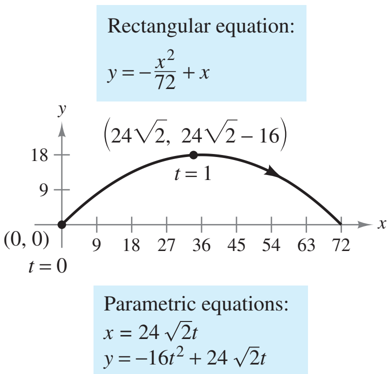
$$x = 24\sqrt{2}\,t \qquad \text{Parametric equation for } x$$

and

$$y = -16t^2 + 24\sqrt{2}\,t. \qquad \text{Parametric equation for } y$$

From this set of equations, you can determine that at time $t = 0$, the object is at the point $(0, 0)$. Similarly, at time $t = 1$, the object is at the point $(24\sqrt{2}, 24\sqrt{2} - 16)$, and so on. (You will learn a method for determining this particular set of parametric equations—the equations of motion—later, in Section 12.3.)

For this particular motion problem, x and y are continuous functions of t, and the resulting path is called a **plane curve.**



Curvilinear motion: two variables for position, one variable for time
Figure 10.19

Animation

Definition of a Plane Curve

If f and g are continuous functions of t on an interval I, then the equations

$$x = f(t) \quad \text{and} \quad y = g(t)$$

are called **parametric equations** and t is called the **parameter.** The set of points (x, y) obtained as t varies over the interval I is called the **graph** of the parametric equations. Taken together, the parametric equations and the graph are called a **plane curve,** denoted by C.

NOTE At times it is important to distinguish between a graph (the set of points) and a curve (the points together with their defining parametric equations). When it is important, we will make the distinction explicit. When it is not important, we will use C to represent the graph or the curve.

When sketching (by hand) a curve represented by a set of parametric equations, you can plot points in the xy-plane. Each set of coordinates (x, y) is determined from a value chosen for the parameter t. By plotting the resulting points in order of increasing values of t, the curve is traced out in a specific direction. This is called the **orientation** of the curve.

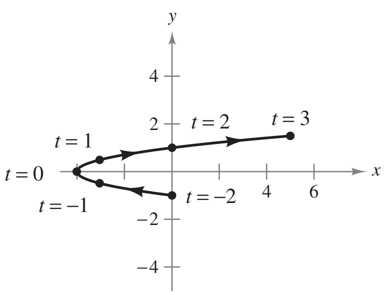
EXAMPLE 1 Sketching a Curve

Sketch the curve described by the parametric equations

$$x = t^2 - 4 \quad \text{and} \quad y = \frac{t}{2}, \quad -2 \le t \le 3.$$

Solution For values of t on the given interval, the parametric equations yield the points (x, y) shown in the table.

t	-2	-1	0	1	2	3
x	0	-3	-4	-3	0	5
y	-1	$-\frac{1}{2}$	0	$\frac{1}{2}$	1	$\frac{3}{2}$



Parametric equations:
$x = t^2 - 4$ and $y = \frac{t}{2}, -2 \le t \le 3$

Figure 10.20

Editable Graph

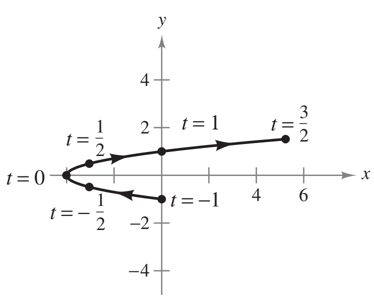
By plotting these points in order of increasing t and using the continuity of f and g, you obtain the curve C shown in Figure 10.20. Note that the arrows on the curve indicate its orientation as t increases from -2 to 3.

Try It Exploration A Exploration B

NOTE From the Vertical Line Test, you can see that the graph shown in Figure 10.20 does not define y as a function of x. This points out one benefit of parametric equations—they can be used to represent graphs that are more general than graphs of functions.

It often happens that two different sets of parametric equations have the same graph. For example, the set of parametric equations

$$x = 4t^2 - 4 \quad \text{and} \quad y = t, \quad -1 \le t \le \frac{3}{2}$$



Parametric equations:
$x = 4t^2 - 4$ and $y = t, -1 \le t \le \frac{3}{2}$

Figure 10.21

has the same graph as the set given in Example 1. However, comparing the values of t in Figures 10.20 and 10.21, you can see that the second graph is traced out more *rapidly* (considering t as time) than the first graph. So, in applications, different parametric representations can be used to represent various *speeds* at which objects travel along a given path.

TECHNOLOGY Most graphing utilities have a *parametric* graphing mode. If you have access to such a utility, use it to confirm the graphs shown in Figures 10.20 and 10.21. Does the curve given by

$$x = 4t^2 - 8t \quad \text{and} \quad y = 1 - t, \quad -\tfrac{1}{2} \le t \le 2$$

represent the same graph as that shown in Figures 10.20 and 10.21? What do you notice about the *orientation* of this curve?

Eliminating the Parameter

Finding a rectangular equation that represents the graph of a set of parametric equations is called **eliminating the parameter.** For instance, you can eliminate the parameter from the set of parametric equations in Example 1 as follows.

Parametric equations	⇨	Solve for t in one equation.	⇨	Substitute into second equation.	⇨	Rectangular equation
$x = t^2 - 4$ $y = t/2$		$t = 2y$		$x = (2y)^2 - 4$		$x = 4y^2 - 4$

Once you have eliminated the parameter, you can recognize that the equation $x = 4y^2 - 4$ represents a parabola with a horizontal axis and vertex at $(-4, 0)$, as shown in Figure 10.20.

The range of x and y implied by the parametric equations may be altered by the change to rectangular form. In such instances the domain of the rectangular equation must be adjusted so that its graph matches the graph of the parametric equations. Such a situation is demonstrated in the next example.

EXAMPLE 2 Adjusting the Domain After Eliminating the Parameter

Sketch the curve represented by the equations

$$x = \frac{1}{\sqrt{t+1}} \quad \text{and} \quad y = \frac{t}{t+1}, \quad t > -1$$

by eliminating the parameter and adjusting the domain of the resulting rectangular equation.

Solution Begin by solving one of the parametric equations for t. For instance, you can solve the first equation for t as follows.

$$x = \frac{1}{\sqrt{t+1}} \qquad \text{Parametric equation for } x$$

$$x^2 = \frac{1}{t+1} \qquad \text{Square each side.}$$

$$t + 1 = \frac{1}{x^2}$$

$$t = \frac{1}{x^2} - 1 = \frac{1 - x^2}{x^2} \qquad \text{Solve for } t.$$

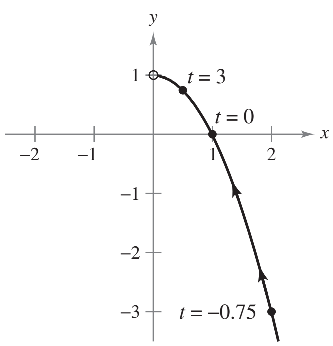
Now, substituting into the parametric equation for y produces

$$y = \frac{t}{t+1} \qquad \text{Parametric equation for } y$$

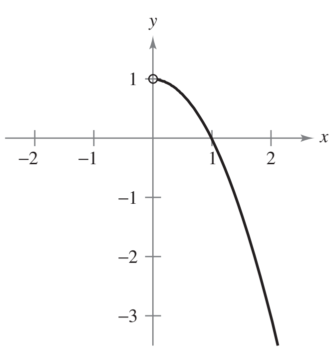
$$y = \frac{(1 - x^2)/x^2}{[(1 - x^2)/x^2] + 1} \qquad \text{Substitute } (1 - x^2)/x^2 \text{ for } t.$$

$$y = 1 - x^2. \qquad \text{Simplify.}$$

The rectangular equation, $y = 1 - x^2$, is defined for all values of x, but from the parametric equation for x you can see that the curve is defined only when $t > -1$. This implies that you should restrict the domain of x to positive values, as shown in Figure 10.22.



Parametric equations:
$x = \dfrac{1}{\sqrt{t+1}}$, $y = \dfrac{t}{t+1}$, $t > -1$



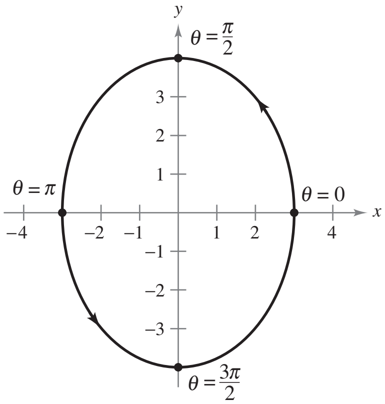
Rectangular equation:
$y = 1 - x^2$, $x > 0$

Figure 10.22

Editable Graph

Try It

Exploration A



Parametric equations:
$x = 3\cos\theta,\ \ y = 4\sin\theta$
Rectangular equation:
$\frac{x^2}{9} + \frac{y^2}{16} = 1$

Figure 10.23

Editable Graph

It is not necessary for the parameter in a set of parametric equations to represent time. The next example uses an *angle* as the parameter.

EXAMPLE 3 Using Trigonometry to Eliminate a Parameter

Sketch the curve represented by

$$x = 3\cos\theta \quad \text{and} \quad y = 4\sin\theta, \quad 0 \le \theta \le 2\pi$$

by eliminating the parameter and finding the corresponding rectangular equation.

Solution Begin by solving for $\cos\theta$ and $\sin\theta$ in the given equations.

$$\cos\theta = \frac{x}{3} \quad \text{and} \quad \sin\theta = \frac{y}{4} \qquad \text{Solve for } \cos\theta \text{ and } \sin\theta.$$

Next, make use of the identity $\sin^2\theta + \cos^2\theta = 1$ to form an equation involving only x and y.

$$\cos^2\theta + \sin^2\theta = 1 \qquad \text{Trigonometric identity}$$

$$\left(\frac{x}{3}\right)^2 + \left(\frac{y}{4}\right)^2 = 1 \qquad \text{Substitute.}$$

$$\frac{x^2}{9} + \frac{y^2}{16} = 1 \qquad \text{Rectangular equation}$$

From this rectangular equation you can see that the graph is an ellipse centered at $(0, 0)$, with vertices at $(0, 4)$ and $(0, -4)$ and minor axis of length $2b = 6$, as shown in Figure 10.23. Note that the ellipse is traced out *counterclockwise* as θ varies from 0 to 2π.

Try It | Exploration A | Open Exploration

Using the technique shown in Example 3, you can conclude that the graph of the parametric equations

$$x = h + a\cos\theta \quad \text{and} \quad y = k + b\sin\theta, \quad 0 \le \theta \le 2\pi$$

is the ellipse (traced counterclockwise) given by

$$\frac{(x-h)^2}{a^2} + \frac{(y-k)^2}{b^2} = 1.$$

The graph of the parametric equations

$$x = h + a\sin\theta \quad \text{and} \quad y = k + b\cos\theta, \quad 0 \le \theta \le 2\pi$$

is also the ellipse (traced clockwise) given by

$$\frac{(x-h)^2}{a^2} + \frac{(y-k)^2}{b^2} = 1.$$

Use a graphing utility in *parametric* mode to graph several ellipses.

In Examples 2 and 3, it is important to realize that eliminating the parameter is primarily an *aid to curve sketching*. If the parametric equations represent the path of a moving object, the graph alone is not sufficient to describe the object's motion. You still need the parametric equations to tell you the *position*, *direction*, and *speed* at a given time.

Finding Parametric Equations

The first three examples in this section illustrate techniques for sketching the graph represented by a set of parametric equations. You will now investigate the reverse problem. How can you determine a set of parametric equations for a given graph or a given physical description? From the discussion following Example 1, you know that such a representation is not unique. This is demonstrated further in the following example, which finds two different parametric representations for a given graph.

EXAMPLE 4 Finding Parametric Equations for a Given Graph

Find a set of parametric equations to represent the graph of $y = 1 - x^2$, using each of the following parameters.

a. $t = x$ **b.** The slope $m = \dfrac{dy}{dx}$ at the point (x, y)

Solution

a. Letting $x = t$ produces the parametric equations

$$x = t \quad \text{and} \quad y = 1 - x^2 = 1 - t^2.$$

b. To write x and y in terms of the parameter m, you can proceed as follows.

$$m = \frac{dy}{dx} = -2x \qquad \text{Differentiate } y = 1 - x^2.$$

$$x = -\frac{m}{2} \qquad \text{Solve for } x.$$

This produces a parametric equation for x. To obtain a parametric equation for y, substitute $-m/2$ for x in the original equation.

$$y = 1 - x^2 \qquad \text{Write original rectangular equation.}$$

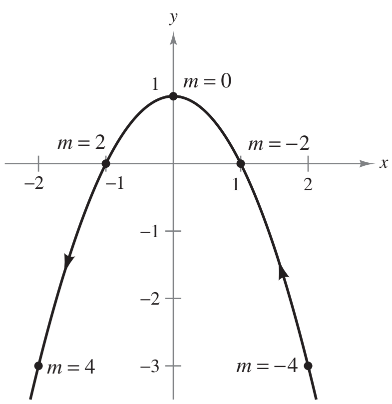
$$y = 1 - \left(-\frac{m}{2}\right)^2 \qquad \text{Substitute } -m/2 \text{ for } x.$$

$$y = 1 - \frac{m^2}{4} \qquad \text{Simplify.}$$

So, the parametric equations are

$$x = -\frac{m}{2} \quad \text{and} \quad y = 1 - \frac{m^2}{4}.$$

In Figure 10.24, note that the resulting curve has a right-to-left orientation as determined by the direction of increasing values of slope m. For part (a), the curve would have the opposite orientation.



Rectangular equation: $y = 1 - x^2$
Parametric equations:
$x = -\frac{m}{2},\ y = 1 - \frac{m^2}{4}$

Figure 10.24

Editable Graph

Try It | Exploration A | Exploration B

TECHNOLOGY To be efficient at using a graphing utility, it is important that you develop skill in representing a graph by a set of parametric equations. The reason for this is that many graphing utilities have only three graphing modes—(1) functions, (2) parametric equations, and (3) polar equations. Most graphing utilities are not programmed to graph a general equation. For instance, suppose you want to graph the hyperbola $x^2 - y^2 = 1$. To graph the hyperbola in *function* mode, you need two equations: $y = \sqrt{x^2 - 1}$ and $y = -\sqrt{x^2 - 1}$. In *parametric* mode, you can represent the graph by $x = \sec t$ and $y = \tan t$.

CYCLOIDS

Galileo first called attention to the cycloid, once recommending that it be used for the arches of bridges. Pascal once spent 8 days attempting to solve many of the problems of cycloids, such as finding the area under one arch, and the volume of the solid of revolution formed by revolving the curve about a line. The cycloid has so many interesting properties and has caused so many quarrels among mathematicians that it has been called "the Helen of geometry" and "the apple of discord."

EXAMPLE 5 Parametric Equations for a Cycloid

Determine the curve traced by a point P on the circumference of a circle of radius a rolling along a straight line in a plane. Such a curve is called a **cycloid.** View the animation to see how a cycloid is drawn.

Animation

Solution Let the parameter θ be the measure of the circle's rotation, and let the point $P = (x, y)$ begin at the origin. When $\theta = 0$, P is at the origin. When $\theta = \pi$, P is at a maximum point $(\pi a, 2a)$. When $\theta = 2\pi$, P is back on the x-axis at $(2\pi a, 0)$. From Figure 10.25, you can see that $\angle APC = 180° - \theta$. So,

$$\sin\theta = \sin(180° - \theta) = \sin(\angle APC) = \frac{AC}{a} = \frac{BD}{a}$$

$$\cos\theta = -\cos(180° - \theta) = -\cos(\angle APC) = \frac{AP}{-a}$$

which implies that

$$AP = -a\cos\theta \quad \text{and} \quad BD = a\sin\theta.$$

Because the circle rolls along the x-axis, you know that $OD = \overparen{PD} = a\theta$. Furthermore, because $BA = DC = a$, you have

$$x = OD - BD = a\theta - a\sin\theta$$
$$y = BA + AP = a - a\cos\theta.$$

So, the parametric equations are

$$x = a(\theta - \sin\theta) \quad \text{and} \quad y = a(1 - \cos\theta).$$

FOR FURTHER INFORMATION For more information on cycloids, see the article "The Geometry of Rolling Curves" by John Bloom and Lee Whitt in *The American Mathematical Monthly*.

MathArticle

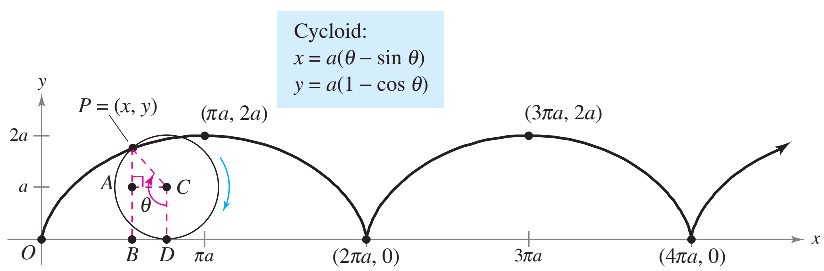


Figure 10.25

Try It   Exploration A

TECHNOLOGY Some graphing utilities allow you to simulate the motion of an object that is moving in the plane or in space. If you have access to such a utility, use it to trace out the path of the cycloid shown in Figure 10.25.

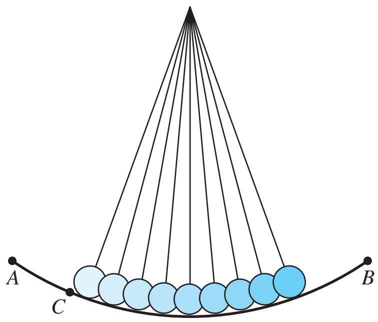
The cycloid in Figure 10.25 has sharp corners at the values $x = 2n\pi a$. Notice that the derivatives $x'(\theta)$ and $y'(\theta)$ are both zero at the points for which $\theta = 2n\pi$.

$$x(\theta) = a(\theta - \sin\theta) \qquad y(\theta) = a(1 - \cos\theta)$$
$$x'(\theta) = a - a\cos\theta \qquad y'(\theta) = a\sin\theta$$
$$x'(2n\pi) = 0 \qquad y'(2n\pi) = 0$$

Between these points, the cycloid is called **smooth.**

Definition of a Smooth Curve

A curve C represented by $x = f(t)$ and $y = g(t)$ on an interval I is called **smooth** if f' and g' are continuous on I and not simultaneously 0, except possibly at the endpoints of I. The curve C is called **piecewise smooth** if it is smooth on each subinterval of some partition of I.



The time required to complete a full swing of the pendulum when starting from point C is only approximately the same as when starting from point A.
Figure 10.26

The Tautochrone and Brachistochrone Problems

The type of curve described in Example 5 is related to one of the most famous pairs of problems in the history of calculus. The first problem (called the **tautochrone problem**) began with Galileo's discovery that the time required to complete a full swing of a given pendulum is *approximately* the same whether it makes a large movement at high speed or a small movement at lower speed (see Figure 10.26). Late in his life, Galileo (1564–1642) realized that he could use this principle to construct a clock. However, he was not able to conquer the mechanics of actual construction. Christian Huygens (1629–1695) was the first to design and construct a working model. In his work with pendulums, Huygens realized that a pendulum does not take exactly the same time to complete swings of varying lengths. (This doesn't affect a pendulum clock, because the length of the circular arc is kept constant by giving the pendulum a slight boost each time it passes its lowest point.) But, in studying the problem, Huygens discovered that a ball rolling back and forth on an inverted cycloid does complete each cycle in exactly the same time.

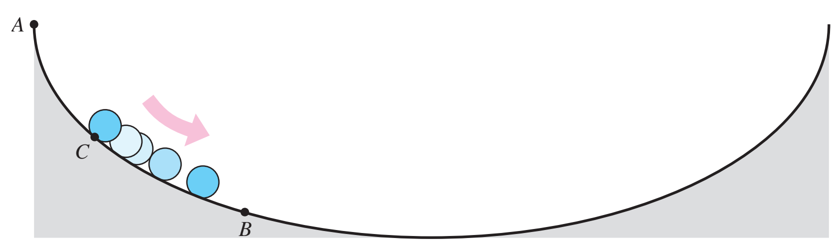
An inverted cycloid is the path down which a ball will roll in the shortest time.
Figure 10.27

JAMES BERNOULLI (1654–1705)

James Bernoulli, also called Jacques, was the older brother of John. He was one of several accomplished mathematicians of the Swiss Bernoulli family. James's mathematical accomplishments have given him a prominent place in the early development of calculus.

MathBio

The second problem, which was posed by John Bernoulli in 1696, is called the **brachistochrone problem**—in Greek, *brachys* means short and chronos means time. The problem was to determine the path down which a particle will slide from point A to point B in the *shortest time*. Several mathematicians took up the challenge, and the following year the problem was solved by Newton, Leibniz, L'Hôpital, John Bernoulli, and James Bernoulli. As it turns out, the solution is not a straight line from A to B, but an inverted cycloid passing through the points A and B, as shown in Figure 10.27. The amazing part of the solution is that a particle starting at rest at *any* other point C of the cycloid between A and B will take exactly the same time to reach B, as shown in Figure 10.28.



A ball starting at point C takes the same time to reach point B as one that starts at point A.
Figure 10.28

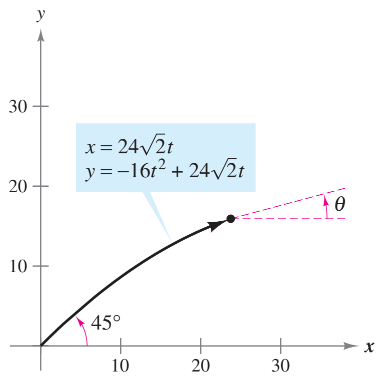
FOR FURTHER INFORMATION To see a proof of the famous brachistochrone problem, see the article "A New Minimization Proof for the Brachistochrone" by Gary Lawlor in *The American Mathematical Monthly*.

MathArticle

Section 10.3 Parametric Equations and Calculus

- **Find the slope of a tangent line to a curve given by a set of parametric equations.**
- **Find the arc length of a curve given by a set of parametric equations.**
- **Find the area of a surface of revolution (parametric form).**

Slope and Tangent Lines



At time t, the angle of elevation of the projectile is θ, the slope of the tangent line at that point.
Figure 10.29

Now that you can represent a graph in the plane by a set of parametric equations, it is natural to ask how to use calculus to study plane curves. To begin, let's take another look at the projectile represented by the parametric equations

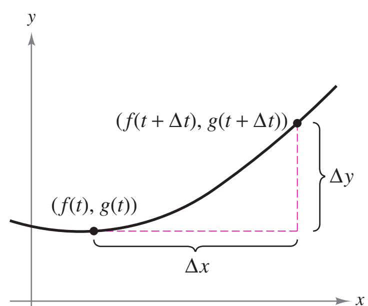
$$x = 24\sqrt{2}t \qquad \text{and} \qquad y = -16t^2 + 24\sqrt{2}t$$

as shown in Figure 10.29. From Section 10.2, you know that these equations enable you to locate the position of the projectile at a given time. You also know that the object is initially projected at an angle of 45°. But how can you find the angle θ representing the object's direction at some other time t? The following theorem answers this question by giving a formula for the slope of the tangent line as a function of t.

THEOREM 10.7 Parametric Form of the Derivative

If a smooth curve C is given by the equations $x = f(t)$ and $y = g(t)$, then the slope of C at (x, y) is

$$\frac{dy}{dx} = \frac{dy/dt}{dx/dt}, \qquad \frac{dx}{dt} \neq 0.$$



The slope of the secant line through the points $(f(t), g(t))$ and $(f(t + \Delta t), g(t + \Delta t))$ is $\Delta y/\Delta x$.
Figure 10.30

Proof In Figure 10.30, consider $\Delta t > 0$ and let

$$\Delta y = g(t + \Delta t) - g(t) \qquad \text{and} \qquad \Delta x = f(t + \Delta t) - f(t).$$

Because $\Delta x \to 0$ as $\Delta t \to 0$, you can write

$$\begin{aligned}\frac{dy}{dx} &= \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x}\\ &= \lim_{\Delta t \to 0} \frac{g(t + \Delta t) - g(t)}{f(t + \Delta t) - f(t)}.\end{aligned}$$

Dividing both the numerator and denominator by Δt, you can use the differentiability of f and g to conclude that

$$\begin{aligned}\frac{dy}{dx} &= \lim_{\Delta t \to 0} \frac{[g(t + \Delta t) - g(t)]/\Delta t}{[f(t + \Delta t) - f(t)]/\Delta t}\\ &= \frac{\displaystyle\lim_{\Delta t \to 0} \frac{g(t + \Delta t) - g(t)}{\Delta t}}{\displaystyle\lim_{\Delta t \to 0} \frac{f(t + \Delta t) - f(t)}{\Delta t}}\\ &= \frac{g'(t)}{f'(t)}\\ &= \frac{dy/dt}{dx/dt}.\end{aligned}$$

STUDY TIP The curve traced out in Example 1 is a circle. Use the formula

$$\frac{dy}{dx} = -\tan t$$

to find the slopes at the points (1, 0) and (0, 1).

EXAMPLE 1 Differentiation and Parametric Form

Find dy/dx for the curve given by $x = \sin t$ and $y = \cos t$.

Solution

$$\frac{dy}{dx} = \frac{dy/dt}{dx/dt} = \frac{-\sin t}{\cos t} = -\tan t$$

Because dy/dx is a function of t, you can use Theorem 10.7 repeatedly to find *higher-order* derivatives. For instance,

$$\frac{d^2y}{dx^2} = \frac{d}{dx}\left[\frac{dy}{dx}\right] = \frac{\frac{d}{dt}\left[\frac{dy}{dx}\right]}{dx/dt}$$ Second derivative

$$\frac{d^3y}{dx^3} = \frac{d}{dx}\left[\frac{d^2y}{dx^2}\right] = \frac{\frac{d}{dt}\left[\frac{d^2y}{dx^2}\right]}{dx/dt}.$$ Third derivative

Try It | Exploration A | Exploration B

The editable graph feature below allows you to edit the graph of a function.

Editable Graph

EXAMPLE 2 Finding Slope and Concavity

For the curve given by

$$x = \sqrt{t} \quad \text{and} \quad y = \frac{1}{4}(t^2 - 4), \quad t \geq 0$$

find the slope and concavity at the point (2, 3).

Solution Because

$$\frac{dy}{dx} = \frac{dy/dt}{dx/dt} = \frac{(1/2)t}{(1/2)t^{-1/2}} = t^{3/2}$$ Parametric form of first derivative

you can find the second derivative to be

$$\frac{d^2y}{dx^2} = \frac{\frac{d}{dt}[dy/dx]}{dx/dt} = \frac{\frac{d}{dt}[t^{3/2}]}{dx/dt} = \frac{(3/2)t^{1/2}}{(1/2)t^{-1/2}} = 3t.$$ Parametric form of second derivative

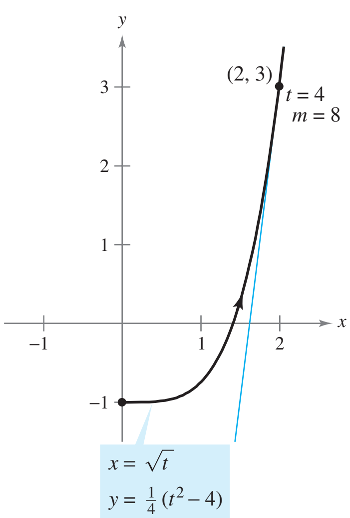
At $(x, y) = (2, 3)$, it follows that $t = 4$, and the slope is

$$\frac{dy}{dx} = (4)^{3/2} = 8.$$

Moreover, when $t = 4$, the second derivative is

$$\frac{d^2y}{dx^2} = 3(4) = 12 > 0$$

and you can conclude that the graph is concave upward at (2, 3), as shown in Figure 10.31.



The graph is concave upward at (2, 3), when $t = 4$.
Figure 10.31

Editable Graph

Try It | Exploration A | Exploration B

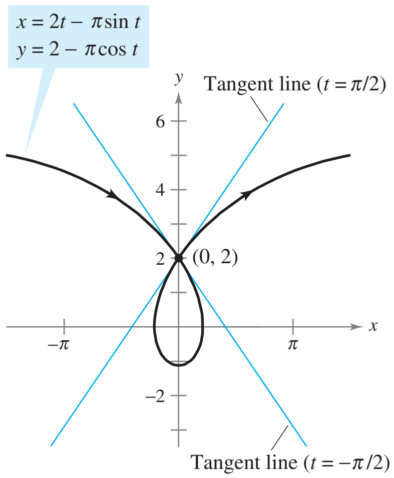
Because the parametric equations $x = f(t)$ and $y = g(t)$ need not define y as a function of x, it is possible for a plane curve to loop around and cross itself. At such points the curve may have more than one tangent line, as shown in the next example.

EXAMPLE 3 A Curve with Two Tangent Lines at a Point

The **prolate cycloid** given by

$$x = 2t - \pi \sin t \qquad \text{and} \qquad y = 2 - \pi \cos t$$

crosses itself at the point $(0, 2)$, as shown in Figure 10.32. Find the equations of both tangent lines at this point.



This prolate cycloid has two tangent lines at the point $(0, 2)$.
Figure 10.32

Editable Graph

Solution Because $x = 0$ and $y = 2$ when $t = \pm\pi/2$, and

$$\frac{dy}{dx} = \frac{dy/dt}{dx/dt} = \frac{\pi \sin t}{2 - \pi \cos t}$$

you have $dy/dx = -\pi/2$ when $t = -\pi/2$ and $dy/dx = \pi/2$ when $t = \pi/2$. So, the two tangent lines at $(0, 2)$ are

$$y - 2 = -\left(\frac{\pi}{2}\right)x \qquad \text{Tangent line when } t = -\frac{\pi}{2}$$

$$y - 2 = \left(\frac{\pi}{2}\right)x. \qquad \text{Tangent line when } t = \frac{\pi}{2}$$

Try It | Exploration A | Exploration B | Open Exploration

If $dy/dt = 0$ and $dx/dt \neq 0$ when $t = t_0$, the curve represented by $x = f(t)$ and $y = g(t)$ has a horizontal tangent at $(f(t_0), g(t_0))$. For instance, in Example 3, the given curve has a horizontal tangent at the point $(0, 2 - \pi)$ (when $t = 0$). Similarly, if $dx/dt = 0$ and $dy/dt \neq 0$ when $t = t_0$, the curve represented by $x = f(t)$ and $y = g(t)$ has a vertical tangent at $(f(t_0), g(t_0))$.

Arc Length

You have seen how parametric equations can be used to describe the path of a particle moving in the plane. You will now develop a formula for determining the *distance* traveled by the particle along its path.

Recall from Section 7.4 that the formula for the arc length of a curve C given by $y = h(x)$ over the interval $[x_0, x_1]$ is

$$\begin{aligned} s &= \int_{x_0}^{x_1} \sqrt{1 + [h'(x)]^2}\, dx \\ &= \int_{x_0}^{x_1} \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\, dx. \end{aligned}$$

If C is represented by the parametric equations $x = f(t)$ and $y = g(t)$, $a \leq t \leq b$, and if $dx/dt = f'(t) > 0$, you can write

$$\begin{aligned} s = \int_{x_0}^{x_1} \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\, dx &= \int_{x_0}^{x_1} \sqrt{1 + \left(\frac{dy/dt}{dx/dt}\right)^2}\, dx \\ &= \int_a^b \sqrt{\frac{(dx/dt)^2 + (dy/dt)^2}{(dx/dt)^2}}\, \frac{dx}{dt}\, dt \\ &= \int_a^b \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}\, dt \\ &= \int_a^b \sqrt{[f'(t)]^2 + [g'(t)]^2}\, dt. \end{aligned}$$

NOTE When applying the arc length formula to a curve, be sure that the curve is traced out only once on the interval of integration. For instance, the circle given by $x = \cos t$ and $y = \sin t$ is traced out once on the interval $0 \le t \le 2\pi$, but is traced out twice on the interval $0 \le t \le 4\pi$.

THEOREM 10.8 Arc Length in Parametric Form

If a smooth curve C is given by $x = f(t)$ and $y = g(t)$ such that C does not intersect itself on the interval $a \le t \le b$ (except possibly at the endpoints), then the arc length of C over the interval is given by

$$s = \int_a^b \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}\, dt = \int_a^b \sqrt{[f'(t)]^2 + [g'(t)]^2}\, dt.$$

In the preceding section you saw that if a circle rolls along a line, a point on its circumference will trace a path called a cycloid. If the circle rolls around the circumference of another circle, the path of the point is an **epicycloid.** The next example shows how to find the arc length of an epicycloid.

Arch of a Cycloid

The arc length of an arch of a cycloid was first calculated in 1658 by British architect and mathematician Christopher Wren, famous for rebuilding many buildings and churches in London, including St. Paul's Cathedral.

EXAMPLE 4 Finding Arc Length

A circle of radius 1 rolls around the circumference of a larger circle of radius 4, as shown in Figure 10.33. The epicycloid traced by a point on the circumference of the smaller circle is given by

$$x = 5\cos t - \cos 5t$$

and

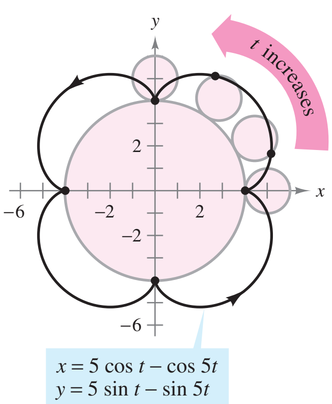
$$y = 5\sin t - \sin 5t.$$

Find the distance traveled by the point in one complete trip about the larger circle.

Solution Before applying Theorem 10.8, note in Figure 10.33 that the curve has sharp points when $t = 0$ and $t = \pi/2$. Between these two points, dx/dt and dy/dt are not simultaneously 0. So, the portion of the curve generated from $t = 0$ to $t = \pi/2$ is smooth. To find the total distance traveled by the point, you can find the arc length of that portion lying in the first quadrant and multiply by 4.

$$\begin{aligned}
s &= 4\int_0^{\pi/2} \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}\, dt && \text{Parametric form for arc length}\\
&= 4\int_0^{\pi/2} \sqrt{(-5\sin t + 5\sin 5t)^2 + (5\cos t - 5\cos 5t)^2}\, dt\\
&= 20\int_0^{\pi/2} \sqrt{2 - 2\sin t\sin 5t - 2\cos t\cos 5t}\, dt\\
&= 20\int_0^{\pi/2} \sqrt{2 - 2\cos 4t}\, dt\\
&= 20\int_0^{\pi/2} \sqrt{4\sin^2 2t}\, dt && \text{Trigonometric identity}\\
&= 40\int_0^{\pi/2} \sin 2t\, dt\\
&= -20\Big[\cos 2t\Big]_0^{\pi/2}\\
&= 40
\end{aligned}$$

For the epicycloid shown in Figure 10.33, an arc length of 40 seems about right because the circumference of a circle of radius 6 is $2\pi r = 12\pi \approx 37.7$.

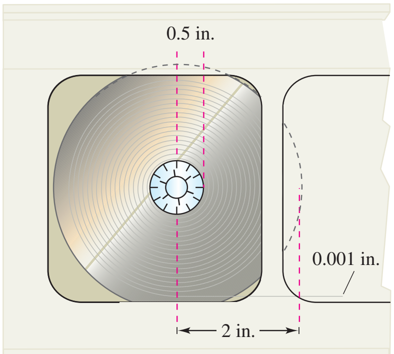


An epicycloid is traced by a point on the smaller circle as it rolls around the larger circle.
Figure 10.33

Editable Graph

Try It

Exploration A



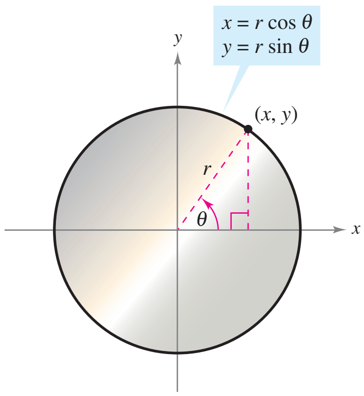


Figure 10.34

EXAMPLE 5 Length of a Recording Tape

A recording tape 0.001 inch thick is wound around a reel whose inner radius is 0.5 inch and whose outer radius is 2 inches, as shown in Figure 10.34. How much tape is required to fill the reel?

Solution To create a model for this problem, assume that as the tape is wound around the reel its distance r from the center increases linearly at a rate of 0.001 inch per revolution, or

$$r = (0.001)\frac{\theta}{2\pi} = \frac{\theta}{2000\pi}, \qquad 1000\pi \le \theta \le 4000\pi$$

where θ is measured in radians. You can determine the coordinates of the point (x, y) corresponding to a given radius to be

$$x = r\cos\theta$$

and

$$y = r\sin\theta.$$

Substituting for r, you obtain the parametric equations

$$x = \left(\frac{\theta}{2000\pi}\right)\cos\theta \quad \text{and} \quad y = \left(\frac{\theta}{2000\pi}\right)\sin\theta.$$

You can use the arc length formula to determine the total length of the tape to be

$$\begin{aligned}
s &= \int_{1000\pi}^{4000\pi} \sqrt{\left(\frac{dx}{d\theta}\right)^2 + \left(\frac{dy}{d\theta}\right)^2}\, d\theta \\
&= \frac{1}{2000\pi}\int_{1000\pi}^{4000\pi} \sqrt{(-\theta\sin\theta + \cos\theta)^2 + (\theta\cos\theta + \sin\theta)^2}\, d\theta \\
&= \frac{1}{2000\pi}\int_{1000\pi}^{4000\pi} \sqrt{\theta^2 + 1}\, d\theta \\
&= \frac{1}{2000\pi}\left(\frac{1}{2}\right)\Big[\theta\sqrt{\theta^2+1} + \ln\left|\theta + \sqrt{\theta^2+1}\right|\Big]_{1000\pi}^{4000\pi} && \text{Integration tables (Appendix B), Formula 26} \\
&\approx 11{,}781 \text{ inches} \\
&\approx 982 \text{ feet}
\end{aligned}$$

Try It **Exploration A**

FOR FURTHER INFORMATION For more information on the mathematics of recording tape, see "Tape Counters" by Richard L. Roth in *The American Mathematical Monthly*.

MathArticle

NOTE The graph of $r = a\theta$ is called the **spiral of Archimedes**. The graph of $r = \theta/2000\pi$ (in Example 5) is of this form.

The length of the tape in Example 5 can be approximated by adding the circumferences of circular pieces of tape. The smallest circle has a radius of 0.501 and the largest has a radius of 2.

$$\begin{aligned}
s &\approx 2\pi(0.501) + 2\pi(0.502) + 2\pi(0.503) + \cdots + 2\pi(2.000) \\
&= \sum_{i=1}^{1500} 2\pi(0.5 + 0.001i) \\
&= 2\pi[1500(0.5) + 0.001(1500)(1501)/2] \\
&\approx 11{,}786 \text{ inches}
\end{aligned}$$

Area of a Surface of Revolution

You can use the formula for the area of a surface of revolution in rectangular form to develop a formula for surface area in parametric form.

THEOREM 10.9 Area of a Surface of Revolution

If a smooth curve C given by $x = f(t)$ and $y = g(t)$ does not cross itself on an interval $a \le t \le b$, then the area S of the surface of revolution formed by revolving C about the coordinate axes is given by the following.

1. $S = 2\pi \int_a^b g(t) \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}\, dt$ Revolution about the x-axis: $g(t) \ge 0$

2. $S = 2\pi \int_a^b f(t) \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}\, dt$ Revolution about the y-axis: $f(t) \ge 0$

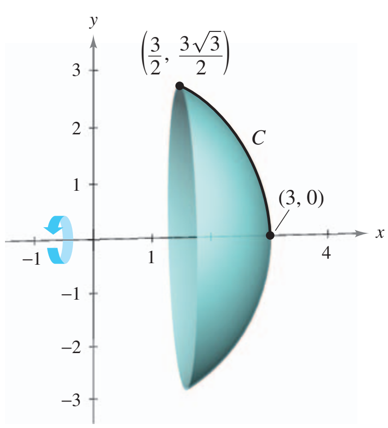
These formulas are easy to remember if you think of the differential of arc length as

$$ds = \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}\, dt.$$

Then the formulas are written as follows.

1. $S = 2\pi \int_a^b g(t)\, ds$ 2. $S = 2\pi \int_a^b f(t)\, ds$

EXAMPLE 6 Finding the Area of a Surface of Revolution



This surface of revolution has a surface area of 9π.
Figure 10.35

Rotatable Graph

Let C be the arc of the circle

$$x^2 + y^2 = 9$$

from $(3, 0)$ to $(3/2, 3\sqrt{3}/2)$, as shown in Figure 10.35. Find the area of the surface formed by revolving C about the x-axis.

Solution You can represent C parametrically by the equations

$$x = 3\cos t \quad \text{and} \quad y = 3\sin t, \quad 0 \le t \le \pi/3.$$

(Note that you can determine the interval for t by observing that $t = 0$ when $x = 3$ and $t = \pi/3$ when $x = 3/2$.) On this interval, C is smooth and y is nonnegative, and you can apply Theorem 10.9 to obtain a surface area of

$$S = 2\pi \int_0^{\pi/3} (3\sin t)\sqrt{(-3\sin t)^2 + (3\cos t)^2}\, dt \qquad \text{Formula for area of a surface of revolution}$$

$$= 6\pi \int_0^{\pi/3} \sin t \sqrt{9(\sin^2 t + \cos^2 t)}\, dt$$

$$= 6\pi \int_0^{\pi/3} 3\sin t\, dt \qquad \text{Trigonometric identity}$$

$$= -18\pi \Big[\cos t\Big]_0^{\pi/3}$$

$$= -18\pi\left(\frac{1}{2} - 1\right)$$

$$= 9\pi.$$

Try It Exploration A

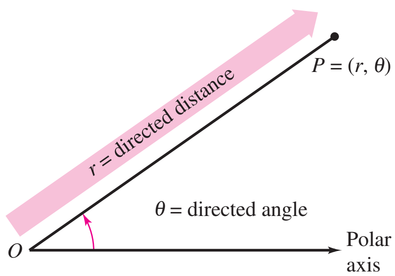
Section 10.4 Polar Coordinates and Polar Graphs

- **Understand the polar coordinate system.**
- **Rewrite rectangular coordinates and equations in polar form and vice versa.**
- **Sketch the graph of an equation given in polar form.**
- **Find the slope of a tangent line to a polar graph.**
- **Identify several types of special polar graphs.**

Polar Coordinates

So far, you have been representing graphs as collections of points (x, y) on the rectangular coordinate system. The corresponding equations for these graphs have been in either rectangular or parametric form. In this section you will study a coordinate system called the **polar coordinate system.**

To form the polar coordinate system in the plane, fix a point O, called the **pole** (or **origin**), and construct from O an initial ray called the **polar axis,** as shown in Figure 10.36. Then each point P in the plane can be assigned **polar coordinates** (r, θ), as follows.

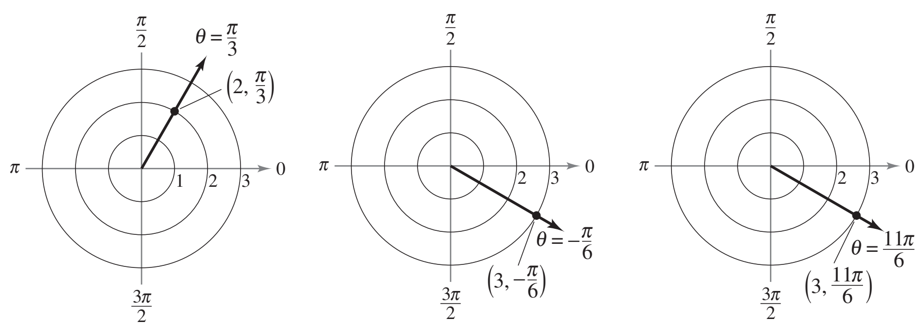


Polar coordinates
Figure 10.36

$r =$ *directed distance* from O to P

$\theta =$ *directed angle*, counterclockwise from polar axis to segment $\overline{OP}$

Figure 10.37 shows three points on the polar coordinate system. Notice that in this system, it is convenient to locate points with respect to a grid of concentric circles intersected by **radial lines** through the pole.



(a) **(b)** **(c)**
Figure 10.37

With rectangular coordinates, each point (x, y) has a unique representation. This is not true with polar coordinates. For instance, the coordinates (r, θ) and $(r, 2\pi + \theta)$ represent the same point [see parts (b) and (c) in Figure 10.37]. Also, because r is a *directed distance*, the coordinates (r, θ) and $(-r, \theta + \pi)$ represent the same point. In general, the point (r, θ) can be written as

$$(r, \theta) = (r, \theta + 2n\pi)$$

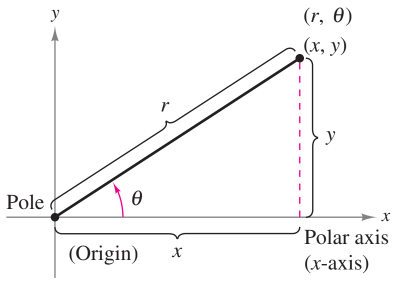
or

$$(r, \theta) = (-r, \theta + (2n + 1)\pi)$$

where n is any integer. Moreover, the pole is represented by $(0, \theta)$, where θ is any angle.

Polar Coordinates

The mathematician credited with first using polar coordinates was James Bernoulli, who introduced them in 1691. However, there is some evidence that it may have been Isaac Newton who first used them.



Relating polar and rectangular coordinates
Figure 10.38

Coordinate Conversion

To establish the relationship between polar and rectangular coordinates, let the polar axis coincide with the positive x-axis and the pole with the origin, as shown in Figure 10.38. Because (x, y) lies on a circle of radius r, it follows that $r^2 = x^2 + y^2$. Moreover, for $r > 0$, the definition of the trigonometric functions implies that

$$\tan\theta = \frac{y}{x}, \qquad \cos\theta = \frac{x}{r}, \qquad \text{and} \qquad \sin\theta = \frac{y}{r}.$$

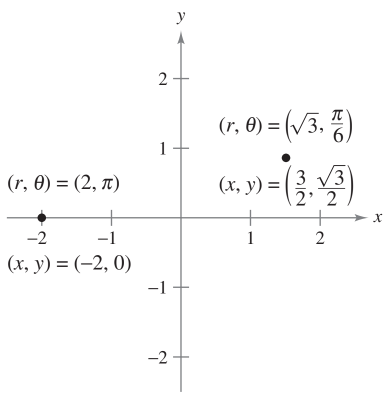
If $r < 0$, you can show that the same relationships hold.

THEOREM 10.10 Coordinate Conversion

The polar coordinates (r, θ) of a point are related to the rectangular coordinates (x, y) of the point as follows.

1. $x = r\cos\theta$ **2.** $\tan\theta = \dfrac{y}{x}$

$y = r\sin\theta$ $r^2 = x^2 + y^2$



To convert from polar to rectangular coordinates, let $x = r\cos\theta$ and $y = r\sin\theta$.
Figure 10.39

EXAMPLE 1 Polar-to-Rectangular Conversion

a. For the point $(r, \theta) = (2, \pi)$,

$$x = r\cos\theta = 2\cos\pi = -2 \qquad \text{and} \qquad y = r\sin\theta = 2\sin\pi = 0.$$

So, the rectangular coordinates are $(x, y) = (-2, 0)$.

b. For the point $(r, \theta) = \left(\sqrt{3}, \pi/6\right)$,

$$x = \sqrt{3}\cos\frac{\pi}{6} = \frac{3}{2} \qquad \text{and} \qquad y = \sqrt{3}\sin\frac{\pi}{6} = \frac{\sqrt{3}}{2}.$$

So, the rectangular coordinates are $(x, y) = \left(3/2, \sqrt{3}/2\right)$.

See Figure 10.39.

Try It   Exploration A   Exploration B

EXAMPLE 2 Rectangular-to-Polar Conversion

a. For the second quadrant point $(x, y) = (-1, 1)$,

$$\tan\theta = \frac{y}{x} = -1 \quad \Rightarrow \quad \theta = \frac{3\pi}{4}.$$

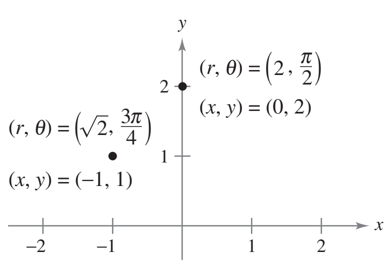
Because θ was chosen to be in the same quadrant as (x, y), you should use a positive value of r.

$$\begin{aligned} r &= \sqrt{x^2 + y^2} \\ &= \sqrt{(-1)^2 + (1)^2} \\ &= \sqrt{2} \end{aligned}$$

This implies that *one* set of polar coordinates is $(r, \theta) = \left(\sqrt{2}, 3\pi/4\right)$.

b. Because the point $(x, y) = (0, 2)$ lies on the positive y-axis, choose $\theta = \pi/2$ and $r = 2$, and one set of polar coordinates is $(r, \theta) = (2, \pi/2)$.

See Figure 10.40.



To convert from rectangular to polar coordinates, let $\tan\theta = y/x$ and $r = \sqrt{x^2 + y^2}$.
Figure 10.40

Try It   Exploration A

Polar Graphs

One way to sketch the graph of a polar equation is to convert to rectangular coordinates and then sketch the graph of the rectangular equation.

EXAMPLE 3 Graphing Polar Equations

Describe the graph of each polar equation. Confirm each description by converting to a rectangular equation.

a. $r = 2$ **b.** $\theta = \dfrac{\pi}{3}$ **c.** $r = \sec\theta$

Solution

a. The graph of the polar equation $r = 2$ consists of all points that are two units from the pole. In other words, this graph is a circle centered at the origin with a radius of 2. [See Figure 10.41(a).] You can confirm this by using the relationship $r^2 = x^2 + y^2$ to obtain the rectangular equation

$$x^2 + y^2 = 2^2. \qquad \text{Rectangular equation}$$

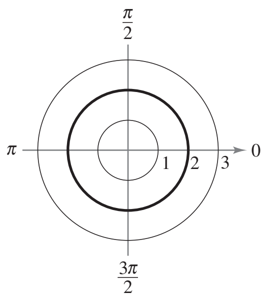
b. The graph of the polar equation $\theta = \pi/3$ consists of all points on the line that makes an angle of $\pi/3$ with the positive x-axis. [See Figure 10.41(b).] You can confirm this by using the relationship $\tan\theta = y/x$ to obtain the rectangular equation

$$y = \sqrt{3}\,x. \qquad \text{Rectangular equation}$$

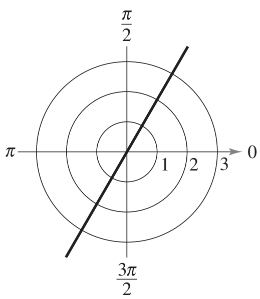
c. The graph of the polar equation $r = \sec\theta$ is not evident by simple inspection, so you can begin by converting to rectangular form using the relationship $r\cos\theta = x$.

$$\begin{aligned} r &= \sec\theta && \text{Polar equation} \\ r\cos\theta &= 1 && \\ x &= 1 && \text{Rectangular equation} \end{aligned}$$

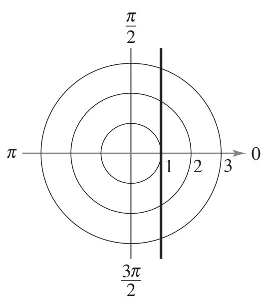
From the rectangular equation, you can see that the graph is a vertical line. [See Figure 10.41(c).]



(a) Circle: $r = 2$



(b) Radial line: $\theta = \dfrac{\pi}{3}$



(c) Vertical line: $r = \sec\theta$

Figure 10.41

Try It | Exploration A

TECHNOLOGY Sketching the graphs of complicated polar equations *by hand* can be tedious. With technology, however, the task is not difficult. If your graphing utility has a *polar* mode, use it to graph the equations in the exercise set. If your graphing utility doesn't have a *polar* mode, but does have a *parametric* mode, you can graph $r = f(\theta)$ by writing the equation as

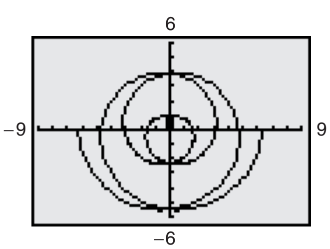
$$\begin{aligned} x &= f(\theta)\cos\theta \\ y &= f(\theta)\sin\theta. \end{aligned}$$

For instance, the graph of $r = \frac{1}{2}\theta$ shown in Figure 10.42 was produced with a graphing calculator in *parametric* mode. This equation was graphed using the parametric equations

$$x = \frac{1}{2}\theta\cos\theta$$

$$y = \frac{1}{2}\theta\sin\theta$$

with the values of θ varying from -4π to 4π. This curve is of the form $r = a\theta$ and is called a **spiral of Archimedes.**



Spiral of Archimedes
Figure 10.42

NOTE One way to sketch the graph of $r = 2\cos 3\theta$ by hand is to make a table of values.

θ	0	$\frac{\pi}{6}$	$\frac{\pi}{3}$	$\frac{\pi}{2}$	$\frac{2\pi}{3}$
r	2	0	-2	0	2

By extending the table and plotting the points, you will obtain the curve shown in Example 4.

EXAMPLE 4 Sketching a Polar Graph

Sketch the graph of $r = 2\cos 3\theta$.

Solution Begin by writing the polar equation in parametric form.

$$x = 2\cos 3\theta \cos\theta \qquad \text{and} \qquad y = 2\cos 3\theta \sin\theta$$

After some experimentation, you will find that the entire curve, which is called a **rose curve,** can be sketched by letting θ vary from 0 to π, as shown in Figure 10.43. If you try duplicating this graph with a graphing utility, you will find that by letting θ vary from 0 to 2π, you will actually trace the entire curve *twice*.

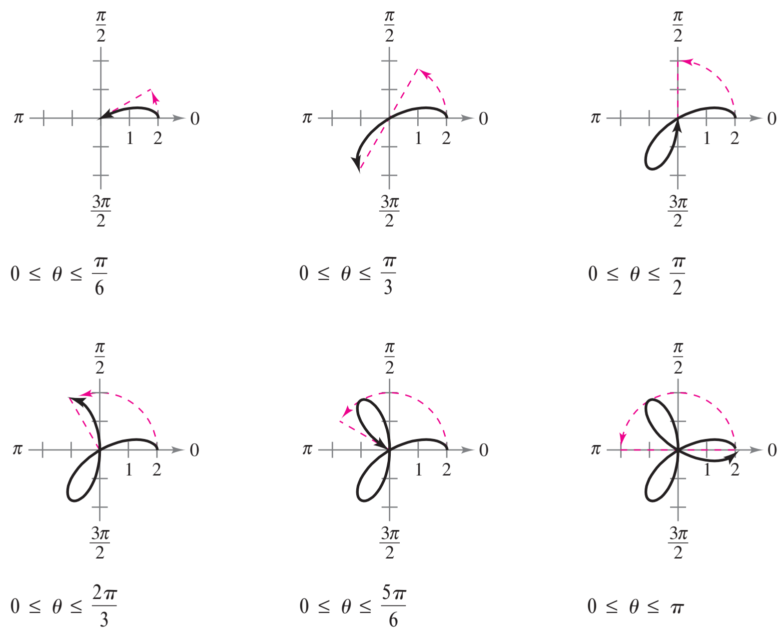
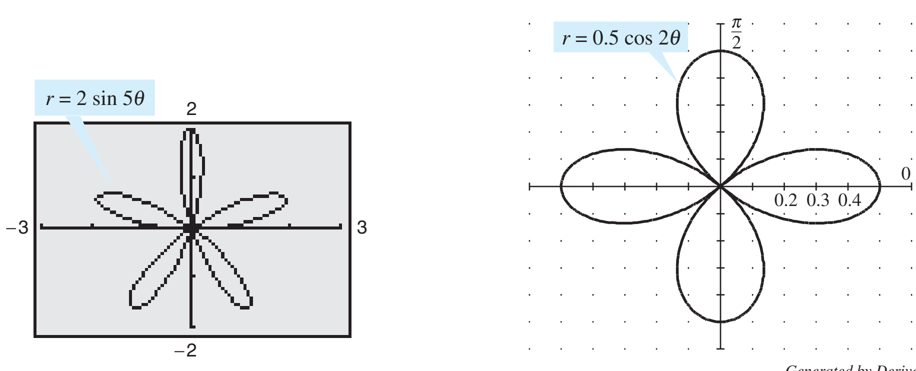


Figure 10.43

Animation

Try It Exploration A Open Exploration

Use a graphing utility to experiment with other rose curves (they are of the form $r = a\cos n\theta$ or $r = a\sin n\theta$). For instance, Figure 10.44 shows the graphs of two other rose curves.



Rose curves
Figure 10.44

Slope and Tangent Lines

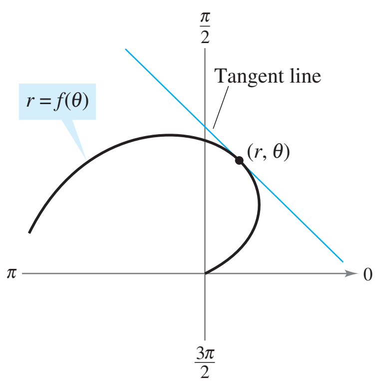
To find the slope of a tangent line to a polar graph, consider a differentiable function given by $r = f(\theta)$. To find the slope in polar form, use the parametric equations

$$x = r\cos\theta = f(\theta)\cos\theta \qquad \text{and} \qquad y = r\sin\theta = f(\theta)\sin\theta.$$

Using the parametric form of dy/dx given in Theorem 10.7, you have

$$\frac{dy}{dx} = \frac{dy/d\theta}{dx/d\theta} = \frac{f(\theta)\cos\theta + f'(\theta)\sin\theta}{-f(\theta)\sin\theta + f'(\theta)\cos\theta}$$

which establishes the following theorem.



Tangent line to polar curve
Figure 10.45

THEOREM 10.11 Slope in Polar Form

If f is a differentiable function of θ, then the *slope* of the tangent line to the graph of $r = f(\theta)$ at the point (r, θ) is

$$\frac{dy}{dx} = \frac{dy/d\theta}{dx/d\theta} = \frac{f(\theta)\cos\theta + f'(\theta)\sin\theta}{-f(\theta)\sin\theta + f'(\theta)\cos\theta}$$

provided that $dx/d\theta \neq 0$ at (r, θ). (See Figure 10.45.)

From Theorem 10.11, you can make the following observations.

1. Solutions to $\dfrac{dy}{d\theta} = 0$ yield horizontal tangents, provided that $\dfrac{dx}{d\theta} \neq 0$.

2. Solutions to $\dfrac{dx}{d\theta} = 0$ yield vertical tangents, provided that $\dfrac{dy}{d\theta} \neq 0$.

If $dy/d\theta$ and $dx/d\theta$ are *simultaneously* 0, no conclusion can be drawn about tangent lines.

EXAMPLE 5 Finding Horizontal and Vertical Tangent Lines

Find the horizontal and vertical tangent lines of $r = \sin\theta$, $0 \leq \theta \leq \pi$.

Solution Begin by writing the equation in parametric form.

$$x = r\cos\theta = \sin\theta\cos\theta$$

and

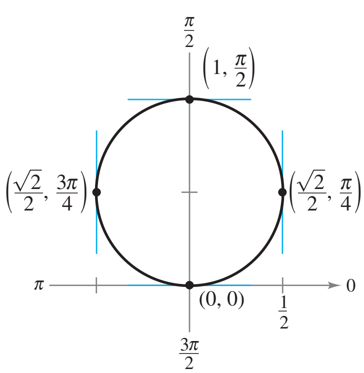
$$y = r\sin\theta = \sin\theta\sin\theta = \sin^2\theta$$

Next, differentiate x and y with respect to θ and set each derivative equal to 0.

$$\frac{dx}{d\theta} = \cos^2\theta - \sin^2\theta = \cos 2\theta = 0 \quad \Longrightarrow \quad \theta = \frac{\pi}{4}, \frac{3\pi}{4}$$

$$\frac{dy}{d\theta} = 2\sin\theta\cos\theta = \sin 2\theta = 0 \quad \Longrightarrow \quad \theta = 0, \frac{\pi}{2}$$

So, the graph has vertical tangent lines at $\left(\sqrt{2}/2, \pi/4\right)$ and $\left(\sqrt{2}/2, 3\pi/4\right)$, and it has horizontal tangent lines at $(0, 0)$ and $(1, \pi/2)$, as shown in Figure 10.46.

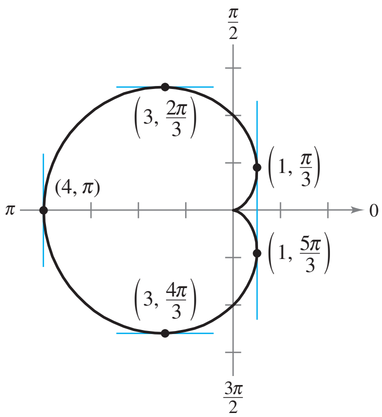


Horizontal and vertical tangent lines of $r = \sin\theta$
Figure 10.46

Editable Graph

Try It

Exploration A



Horizontal and vertical tangent lines of $r = 2(1 - \cos\theta)$
Figure 10.47

Editable Graph

EXAMPLE 6 Finding Horizontal and Vertical Tangent Lines

Find the horizontal and vertical tangents to the graph of $r = 2(1 - \cos\theta)$.

Solution Using $y = r\sin\theta$, differentiate and set $dy/d\theta$ equal to 0.

$$y = r\sin\theta = 2(1 - \cos\theta)\sin\theta$$

$$\frac{dy}{d\theta} = 2[(1 - \cos\theta)(\cos\theta) + \sin\theta(\sin\theta)]$$

$$= -2(2\cos\theta + 1)(\cos\theta - 1) = 0$$

So, $\cos\theta = -\frac{1}{2}$ and $\cos\theta = 1$, and you can conclude that $dy/d\theta = 0$ when $\theta = 2\pi/3, 4\pi/3$, and 0. Similarly, using $x = r\cos\theta$, you have

$$x = r\cos\theta = 2\cos\theta - 2\cos^2\theta$$

$$\frac{dx}{d\theta} = -2\sin\theta + 4\cos\theta\sin\theta = 2\sin\theta(2\cos\theta - 1) = 0.$$

So, $\sin\theta = 0$ or $\cos\theta = \frac{1}{2}$, and you can conclude that $dx/d\theta = 0$ when $\theta = 0, \pi, \pi/3$, and $5\pi/3$. From these results, and from the graph shown in Figure 10.47, you can conclude that the graph has horizontal tangents at $(3, 2\pi/3)$ and $(3, 4\pi/3)$, and has vertical tangents at $(1, \pi/3)$, $(1, 5\pi/3)$, and $(4, \pi)$. This graph is called a **cardioid**. Note that both derivatives $(dy/d\theta$ and $dx/d\theta)$ are 0 when $\theta = 0$. Using this information alone, you don't know whether the graph has a horizontal or vertical tangent line at the pole. From Figure 10.47, however, you can see that the graph has a cusp at the pole.

Try It Exploration A

Theorem 10.11 has an important consequence. Suppose the graph of $r = f(\theta)$ passes through the pole when $\theta = \alpha$ and $f'(\alpha) \neq 0$. Then the formula for dy/dx simplifies as follows.

$$\frac{dy}{dx} = \frac{f'(\alpha)\sin\alpha + f(\alpha)\cos\alpha}{f'(\alpha)\cos\alpha - f(\alpha)\sin\alpha} = \frac{f'(\alpha)\sin\alpha + 0}{f'(\alpha)\cos\alpha - 0} = \frac{\sin\alpha}{\cos\alpha} = \tan\alpha$$

So, the line $\theta = \alpha$ is tangent to the graph at the pole, $(0, \alpha)$.

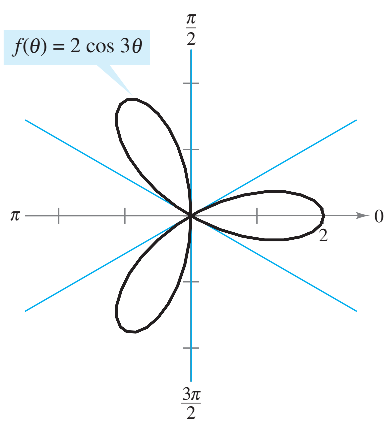
THEOREM 10.12 Tangent Lines at the Pole

If $f(\alpha) = 0$ and $f'(\alpha) \neq 0$, then the line $\theta = \alpha$ is tangent at the pole to the graph of $r = f(\theta)$.

Theorem 10.12 is useful because it states that the zeros of $r = f(\theta)$ can be used to find the tangent lines at the pole. Note that because a polar curve can cross the pole more than once, it can have more than one tangent line at the pole. For example, the rose curve

$$f(\theta) = 2\cos 3\theta$$

has three tangent lines at the pole, as shown in Figure 10.48. For this curve, $f(\theta) = 2\cos 3\theta$ is 0 when θ is $\pi/6$, $\pi/2$, and $5\pi/6$. Moreover, the derivative $f'(\theta) = -6\sin 3\theta$ is not 0 for these values of θ.



This rose curve has three tangent lines $(\theta = \pi/6, \theta = \pi/2$, and $\theta = 5\pi/6)$ at the pole.
Figure 10.48

Special Polar Graphs

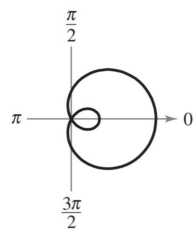
Several important types of graphs have equations that are simpler in polar form than in rectangular form. For example, the polar equation of a circle having a radius of a and centered at the origin is simply $r = a$. Later in the text you will come to appreciate this benefit. For now, several other types of graphs that have simpler equations in polar form are shown below. (Conics are considered in Section 10.6.)

Limaçons

$r = a \pm b\cos\theta$

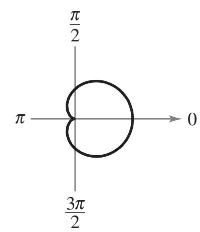
$r = a \pm b\sin\theta$

$(a > 0, b > 0)$



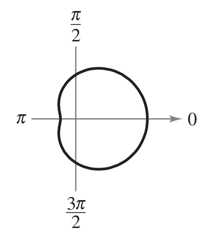
$\frac{a}{b} < 1$

Limaçon with inner loop



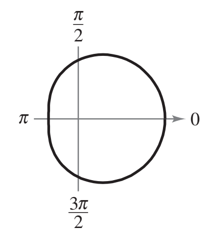
$\frac{a}{b} = 1$

Cardioid (heart-shaped)



$1 < \frac{a}{b} < 2$

Dimpled limaçon



$\frac{a}{b} \geq 2$

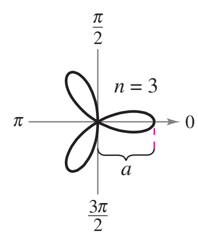
Convex limaçon

Rose Curves

n petals if n is odd

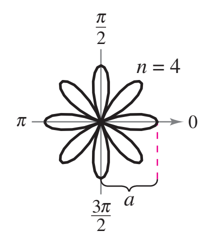
$2n$ petals if n is even

$(n \geq 2)$



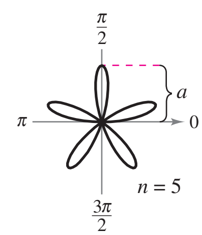
$r = a\cos n\theta$

Rose curve



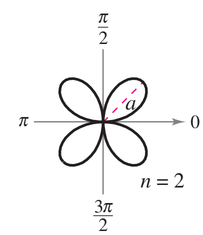
$r = a\cos n\theta$

Rose curve



$r = a\sin n\theta$

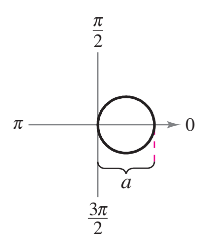
Rose curve



$r = a\sin n\theta$

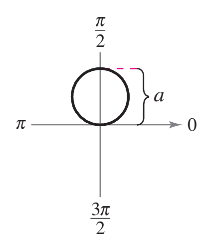
Rose curve

Circles and Lemniscates



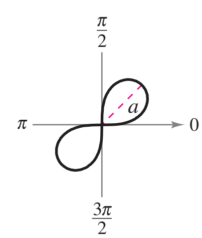
$r = a\cos\theta$

Circle



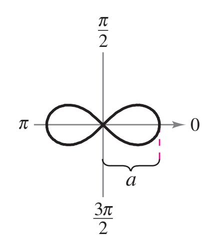
$r = a\sin\theta$

Circle



$r^2 = a^2\sin 2\theta$

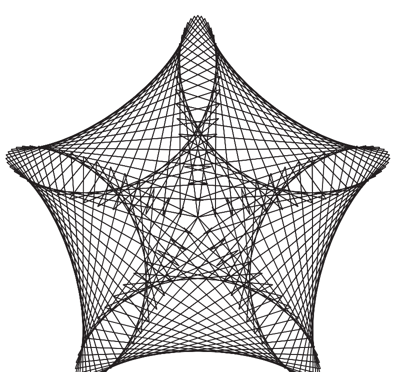
Lemniscate



$r^2 = a^2\cos 2\theta$

Lemniscate

TECHNOLOGY The rose curves described above are of the form $r = a\cos n\theta$ or $r = a\sin n\theta$, where n is a positive integer that is greater than or equal to 2. Use a graphing utility to graph $r = a\cos n\theta$ or $r = a\sin n\theta$ for some noninteger values of n. Are these graphs also rose curves? For example, try sketching the graph of $r = \cos\frac{2}{3}\theta,\ 0 \leq \theta \leq 6\pi$.



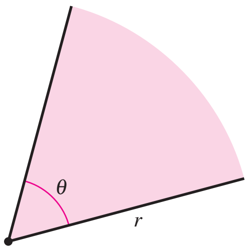
Generated by Maple

FOR FURTHER INFORMATION For more information on rose curves and related curves, see the article "A Rose is a Rose . . ." by Peter M. Maurer in *The American Mathematical Monthly*. The computer-generated graph at the left is the result of an algorithm that Maurer calls "The Rose."

MathArticle

Section 10.5 Area and Arc Length in Polar Coordinates

- **Find the area of a region bounded by a polar graph.**
- **Find the points of intersection of two polar graphs.**
- **Find the arc length of a polar graph.**
- **Find the area of a surface of revolution (polar form).**



The area of a sector of a circle is $A = \frac{1}{2}\theta r^2$.
Figure 10.49

Area of a Polar Region

The development of a formula for the area of a polar region parallels that for the area of a region on the rectangular coordinate system, but uses sectors of a circle instead of rectangles as the basic element of area. In Figure 10.49, note that the area of a circular sector of radius r is given by $\frac{1}{2}\theta r^2$, provided θ is measured in radians.

Consider the function given by $r = f(\theta)$, where f is continuous and nonnegative in the interval given by $\alpha \le \theta \le \beta$. The region bounded by the graph of f and the radial lines $\theta = \alpha$ and $\theta = \beta$ is shown in Figure 10.50(a). To find the area of this region, partition the interval $[\alpha, \beta]$ into n equal subintervals

$$\alpha = \theta_0 < \theta_1 < \theta_2 < \cdot \cdot \cdot < \theta_{n-1} < \theta_n = \beta.$$

Then, approximate the area of the region by the sum of the areas of the n sectors, as shown in Figure 10.50(b).

$$\text{Radius of } i\text{th sector} = f(\theta_i)$$

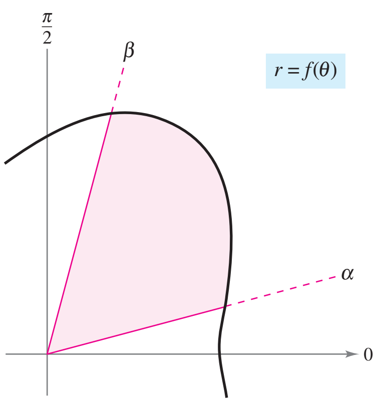
$$\text{Central angle of } i\text{th sector} = \frac{\beta - \alpha}{n} = \Delta\theta$$

$$A \approx \sum_{i=1}^{n} \left(\frac{1}{2}\right) \Delta\theta [f(\theta_i)]^2$$

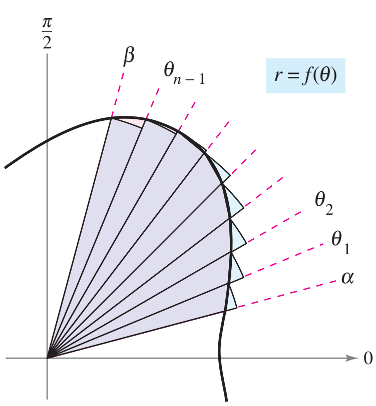
Taking the limit as $n \to \infty$ produces

$$\begin{aligned} A &= \lim_{n\to\infty} \frac{1}{2} \sum_{i=1}^{n} [f(\theta_i)]^2 \, \Delta\theta \\ &= \frac{1}{2} \int_{\alpha}^{\beta} [f(\theta)]^2 \, d\theta \end{aligned}$$

which leads to the following theorem.



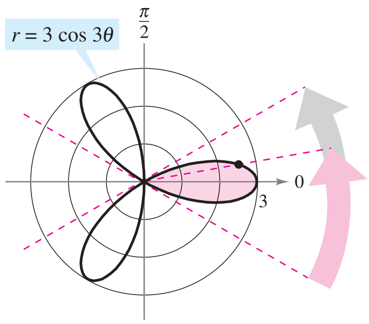
(a)



(b)
Figure 10.50

> **THEOREM 10.13 Area in Polar Coordinates**
>
> If f is continuous and nonnegative on the interval $[\alpha, \beta]$, $0 < \beta - \alpha \le 2\pi$, then the area of the region bounded by the graph of $r = f(\theta)$ between the radial lines $\theta = \alpha$ and $\theta = \beta$ is given by
>
> $$\begin{aligned} A &= \frac{1}{2} \int_{\alpha}^{\beta} [f(\theta)]^2 \, d\theta \\ &= \frac{1}{2} \int_{\alpha}^{\beta} r^2 \, d\theta. \qquad 0 < \beta - \alpha \le 2\pi \end{aligned}$$

NOTE You can use the same formula to find the area of a region bounded by the graph of a continuous *nonpositive* function. However, the formula is not necessarily valid if f takes on both positive *and* negative values in the interval $[\alpha, \beta]$.



The area of one petal of the rose curve that lies between the radial lines $\theta = -\pi/6$ and $\theta = \pi/6$ is $3\pi/4$.
Figure 10.51

Editable Graph

EXAMPLE 1 Finding the Area of a Polar Region

Find the area of one petal of the rose curve given by $r = 3\cos 3\theta$.

Solution In Figure 10.51, you can see that the right petal is traced as θ increases from $-\pi/6$ to $\pi/6$. So, the area is

$$\begin{aligned}
A &= \frac{1}{2}\int_{\alpha}^{\beta} r^2\, d\theta = \frac{1}{2}\int_{-\pi/6}^{\pi/6} (3\cos 3\theta)^2\, d\theta && \text{Formula for area in polar coordinates}\\
&= \frac{9}{2}\int_{-\pi/6}^{\pi/6} \frac{1+\cos 6\theta}{2}\, d\theta && \text{Trigonometric identity}\\
&= \frac{9}{4}\left[\theta + \frac{\sin 6\theta}{6}\right]_{-\pi/6}^{\pi/6}\\
&= \frac{9}{4}\left(\frac{\pi}{6} + \frac{\pi}{6}\right)\\
&= \frac{3\pi}{4}.
\end{aligned}$$

Try It | Exploration A | Open Exploration

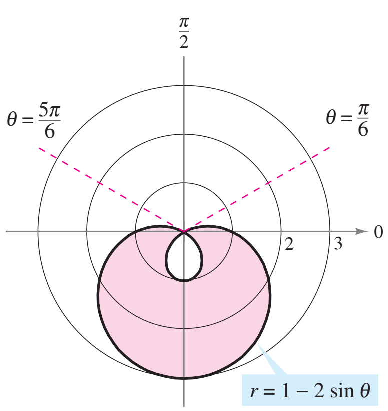
NOTE To find the area of the region lying inside all three petals of the rose curve in Example 1, you could not simply integrate between 0 and 2π. In doing this you would obtain $9\pi/2$, which is twice the area of the three petals. The duplication occurs because the rose curve is traced twice as θ increases from 0 to 2π.

EXAMPLE 2 Finding the Area Bounded by a Single Curve

Find the area of the region lying between the inner and outer loops of the limaçon $r = 1 - 2\sin\theta$.

Solution In Figure 10.52, note that the inner loop is traced as θ increases from $\pi/6$ to $5\pi/6$. So, the area inside the *inner loop* is

$$\begin{aligned}
A_1 &= \frac{1}{2}\int_{\alpha}^{\beta} r^2\, d\theta = \frac{1}{2}\int_{\pi/6}^{5\pi/6} (1 - 2\sin\theta)^2\, d\theta && \text{Formula for area in polar coordinates}\\
&= \frac{1}{2}\int_{\pi/6}^{5\pi/6} (1 - 4\sin\theta + 4\sin^2\theta)\, d\theta\\
&= \frac{1}{2}\int_{\pi/6}^{5\pi/6} \left[1 - 4\sin\theta + 4\left(\frac{1-\cos 2\theta}{2}\right)\right] d\theta && \text{Trigonometric identity}\\
&= \frac{1}{2}\int_{\pi/6}^{5\pi/6} (3 - 4\sin\theta - 2\cos 2\theta)\, d\theta && \text{Simplify.}\\
&= \frac{1}{2}\Big[3\theta + 4\cos\theta - \sin 2\theta\Big]_{\pi/6}^{5\pi/6}\\
&= \frac{1}{2}\left(2\pi - 3\sqrt{3}\right)\\
&= \pi - \frac{3\sqrt{3}}{2}.
\end{aligned}$$



The area between the inner and outer loops is approximately 8.34.
Figure 10.52

Editable Graph

In a similar way, you can integrate from $5\pi/6$ to $13\pi/6$ to find that the area of the region lying inside the outer loop is $A_2 = 2\pi + \left(3\sqrt{3}/2\right)$. The area of the region lying between the two loops is the difference of A_2 and A_1.

$$A = A_2 - A_1 = \left(2\pi + \frac{3\sqrt{3}}{2}\right) - \left(\pi - \frac{3\sqrt{3}}{2}\right) = \pi + 3\sqrt{3} \approx 8.34$$

Try It | Exploration A

Points of Intersection of Polar Graphs

Because a point may be represented in different ways in polar coordinates, care must be taken in determining the points of intersection of two polar graphs. For example, consider the points of intersection of the graphs of

$$r = 1 - 2\cos\theta \quad \text{and} \quad r = 1$$

as shown in Figure 10.53. If, as with rectangular equations, you attempted to find the points of intersection by solving the two equations simultaneously, you would obtain

$$\begin{aligned} r &= 1 - 2\cos\theta && \text{First equation} \\ 1 &= 1 - 2\cos\theta && \text{Substitute } r = 1 \text{ from 2nd equation into 1st equation.} \\ \cos\theta &= 0 && \text{Simplify.} \\ \theta &= \frac{\pi}{2}, \frac{3\pi}{2}. && \text{Solve for } \theta. \end{aligned}$$

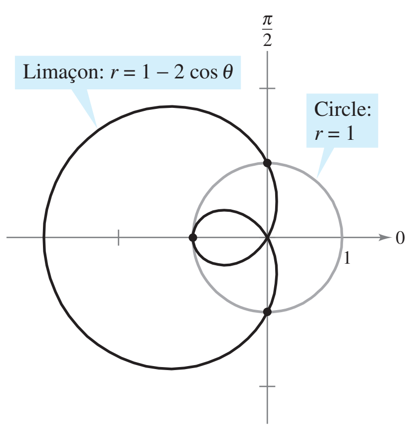
FOR FURTHER INFORMATION For more information on using technology to find points of intersection, see the article "Finding Points of Intersection of Polar-Coordinate Graphs" by Warren W. Esty in *Mathematics Teacher*.

MathArticle

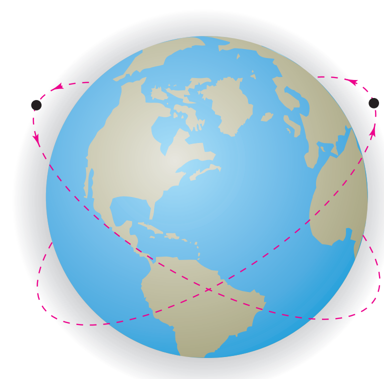
The corresponding points of intersection are $(1, \pi/2)$ and $(1, 3\pi/2)$. However, from Figure 10.53 you can see that there is a *third* point of intersection that did not show up when the two polar equations were solved simultaneously. (This is one reason why you should sketch a graph when finding the area of a polar region.) The reason the third point was not found is that it does not occur with the same coordinates in the two graphs. On the graph of $r = 1$, the point occurs with coordinates $(1, \pi)$, but on the graph of $r = 1 - 2\cos\theta$, the point occurs with coordinates $(-1, 0)$.

You can compare the problem of finding points of intersection of two polar graphs with that of finding collision points of two satellites in intersecting orbits about Earth, as shown in Figure 10.54. The satellites will not collide as long as they reach the points of intersection at different times (θ-values). Collisions will occur only at the points of intersection that are "simultaneous points"—those reached at the same time (θ-value).

NOTE Because the pole can be represented by $(0, \theta)$, where θ is *any* angle, you should check separately for the pole when finding points of intersection.



Three points of intersection: $(1, \pi/2)$, $(-1, 0)$, $(1, 3\pi/2)$
Figure 10.53



The paths of satellites can cross without causing a collision.
Figure 10.54

Animation

EXAMPLE 3 Finding the Area of a Region Between Two Curves

Find the area of the region common to the two regions bounded by the following curves.

$$r = -6 \cos \theta \qquad \text{Circle}$$

$$r = 2 - 2 \cos \theta \qquad \text{Cardioid}$$

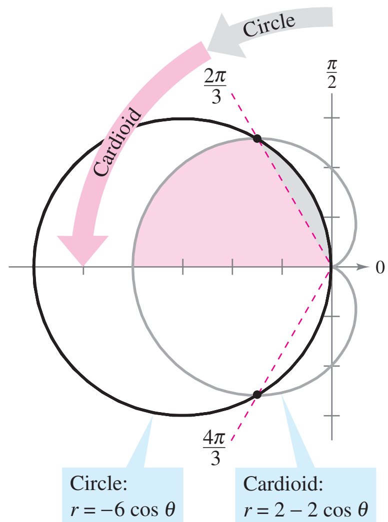


Figure 10.55

Editable Graph

Solution Because both curves are symmetric with respect to the x-axis, you can work with the upper half-plane, as shown in Figure 10.55. The gray shaded region lies between the circle and the radial line $\theta = 2\pi/3$. Because the circle has coordinates $(0, \pi/2)$ at the pole, you can integrate between $\pi/2$ and $2\pi/3$ to obtain the area of this region. The region that is shaded red is bounded by the radial lines $\theta = 2\pi/3$ and $\theta = \pi$ and the cardioid. So, you can find the area of this second region by integrating between $2\pi/3$ and π. The sum of these two integrals gives the area of the common region lying *above* the radial line $\theta = \pi$.

$$\frac{A}{2} = \overbrace{\frac{1}{2}\int_{\pi/2}^{2\pi/3} (-6 \cos \theta)^2 \, d\theta}^{\text{Region between circle and radial line } \theta = 2\pi/3} + \overbrace{\frac{1}{2}\int_{2\pi/3}^{\pi} (2 - 2 \cos \theta)^2 \, d\theta}^{\text{Region between cardioid and radial lines } \theta = 2\pi/3 \text{ and } \theta = \pi}$$

$$= 18 \int_{\pi/2}^{2\pi/3} \cos^2 \theta \, d\theta + \frac{1}{2}\int_{2\pi/3}^{\pi} (4 - 8 \cos \theta + 4 \cos^2 \theta) \, d\theta$$

$$= 9 \int_{\pi/2}^{2\pi/3} (1 + \cos 2\theta) \, d\theta + \int_{2\pi/3}^{\pi} (3 - 4 \cos \theta + \cos 2\theta) \, d\theta$$

$$= 9\left[\theta + \frac{\sin 2\theta}{2}\right]_{\pi/2}^{2\pi/3} + \left[3\theta - 4 \sin \theta + \frac{\sin 2\theta}{2}\right]_{2\pi/3}^{\pi}$$

$$= 9\left(\frac{2\pi}{3} - \frac{\sqrt{3}}{4} - \frac{\pi}{2}\right) + \left(3\pi - 2\pi + 2\sqrt{3} + \frac{\sqrt{3}}{4}\right)$$

$$= \frac{5\pi}{2}$$

$$\approx 7.85$$

Finally, multiplying by 2, you can conclude that the total area is 5π.

Try It Exploration A Exploration B

NOTE To check the reasonableness of the result obtained in Example 3, note that the area of the circular region is $\pi r^2 = 9\pi$. So, it seems reasonable that the area of the region lying inside the circle and the cardioid is 5π.

To see the benefit of polar coordinates for finding the area in Example 3, consider the following integral, which gives the comparable area in rectangular coordinates.

$$\frac{A}{2} = \int_{-4}^{-3/2} \sqrt{2\sqrt{1 - 2x} - x^2 - 2x + 2} \, dx + \int_{-3/2}^{0} \sqrt{-x^2 - 6x} \, dx$$

Use the integration capabilities of a graphing utility to show that you obtain the same area as that found in Example 3.

Arc Length in Polar Form

NOTE When applying the arc length formula to a polar curve, be sure that the curve is traced out only once on the interval of integration. For instance, the rose curve given by $r = \cos 3\theta$ is traced out once on the interval $0 \le \theta \le \pi$, but is traced out twice on the interval $0 \le \theta \le 2\pi$.

The formula for the length of a polar arc can be obtained from the arc length formula for a curve described by parametric equations. (See Exercise 77.)

THEOREM 10.14 Arc Length of a Polar Curve

Let f be a function whose derivative is continuous on an interval $\alpha \le \theta \le \beta$. The length of the graph of $r = f(\theta)$ from $\theta = \alpha$ to $\theta = \beta$ is

$$s = \int_\alpha^\beta \sqrt{[f(\theta)]^2 + [f'(\theta)]^2}\, d\theta = \int_\alpha^\beta \sqrt{r^2 + \left(\frac{dr}{d\theta}\right)^2}\, d\theta.$$

EXAMPLE 4 Finding the Length of a Polar Curve

Find the length of the arc from $\theta = 0$ to $\theta = 2\pi$ for the cardioid

$$r = f(\theta) = 2 - 2\cos\theta$$

as shown in Figure 10.56.

Solution Because $f'(\theta) = 2\sin\theta$, you can find the arc length as follows.

$$\begin{aligned}
s &= \int_\alpha^\beta \sqrt{[f(\theta)]^2 + [f'(\theta)]^2}\, d\theta && \text{Formula for arc length of a polar curve}\\
&= \int_0^{2\pi} \sqrt{(2 - 2\cos\theta)^2 + (2\sin\theta)^2}\, d\theta\\
&= 2\sqrt{2}\int_0^{2\pi} \sqrt{1 - \cos\theta}\, d\theta && \text{Simplify.}\\
&= 2\sqrt{2}\int_0^{2\pi} \sqrt{2\sin^2\frac{\theta}{2}}\, d\theta && \text{Trigonometric identity}\\
&= 4\int_0^{2\pi} \sin\frac{\theta}{2}\, d\theta && \sin\frac{\theta}{2} \ge 0 \text{ for } 0 \le \theta \le 2\pi\\
&= 8\Big[-\cos\frac{\theta}{2}\Big]_0^{2\pi}\\
&= 8(1 + 1)\\
&= 16
\end{aligned}$$

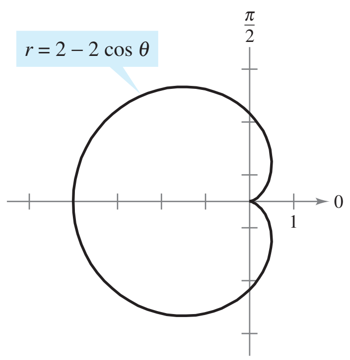


Figure 10.56

Editable Graph

In the fifth step of the solution, it is legitimate to write

$$\sqrt{2\sin^2(\theta/2)} = \sqrt{2}\sin(\theta/2)$$

rather than

$$\sqrt{2\sin^2(\theta/2)} = \sqrt{2}\,|\sin(\theta/2)|$$

because $\sin(\theta/2) \ge 0$ for $0 \le \theta \le 2\pi$.

Try It Exploration A

NOTE Using Figure 10.56, you can determine the reasonableness of this answer by comparing it with the circumference of a circle. For example, a circle of radius $\frac{5}{2}$ has a circumference of $5\pi \approx 15.7$.

Area of a Surface of Revolution

The polar coordinate versions of the formulas for the area of a surface of revolution can be obtained from the parametric versions given in Theorem 10.9, using the equations $x = r\cos\theta$ and $y = r\sin\theta$.

NOTE When using Theorem 10.15, check to see that the graph of $r = f(\theta)$ is traced only once on the interval $\alpha \le \theta \le \beta$. For example, the circle given by $r = \cos\theta$ is traced only once on the interval $0 \le \theta \le \pi$.

THEOREM 10.15 Area of a Surface of Revolution

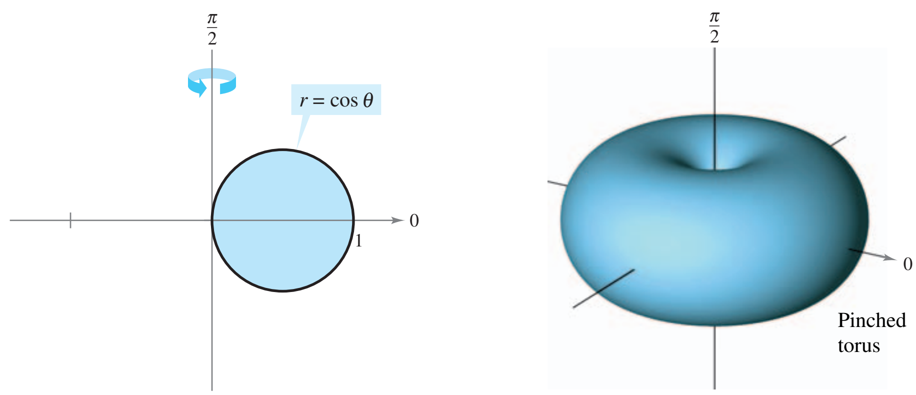
Let f be a function whose derivative is continuous on an interval $\alpha \le \theta \le \beta$. The area of the surface formed by revolving the graph of $r = f(\theta)$ from $\theta = \alpha$ to $\theta = \beta$ about the indicated line is as follows.

1. $S = 2\pi \displaystyle\int_\alpha^\beta f(\theta)\sin\theta\sqrt{[f(\theta)]^2 + [f'(\theta)]^2}\,d\theta$ About the polar axis

2. $S = 2\pi \displaystyle\int_\alpha^\beta f(\theta)\cos\theta\sqrt{[f(\theta)]^2 + [f'(\theta)]^2}\,d\theta$ About the line $\theta = \dfrac{\pi}{2}$

EXAMPLE 5 Finding the Area of a Surface of Revolution

Find the area of the surface formed by revolving the circle $r = f(\theta) = \cos\theta$ about the line $\theta = \pi/2$, as shown in Figure 10.57.



(a) (b)

Figure 10.57

Rotatable Graph

Solution You can use the second formula given in Theorem 10.15 with $f'(\theta) = -\sin\theta$. Because the circle is traced once as θ increases from 0 to π, you have

$$S = 2\pi \int_\alpha^\beta f(\theta)\cos\theta\sqrt{[f(\theta)]^2 + [f'(\theta)]^2}\,d\theta \qquad \text{Formula for area of a surface of revolution}$$

$$= 2\pi \int_0^\pi \cos\theta(\cos\theta)\sqrt{\cos^2\theta + \sin^2\theta}\,d\theta$$

$$= 2\pi \int_0^\pi \cos^2\theta\,d\theta \qquad \text{Trigonometric identity}$$

$$= \pi \int_0^\pi (1 + \cos 2\theta)\,d\theta \qquad \text{Trigonometric identity}$$

$$= \pi\left[\theta + \frac{\sin 2\theta}{2}\right]_0^\pi = \pi^2.$$

Try It Exploration A

Section 10.6 Polar Equations of Conics and Kepler's Laws

- **Analyze and write polar equations of conics.**
- **Understand and use Kepler's Laws of planetary motion.**

EXPLORATION

Graphing Conics Set a graphing utility to *polar* mode and enter polar equations of the form

$$r = \frac{a}{1 \pm b \cos \theta}$$

or

$$r = \frac{a}{1 \pm b \sin \theta}.$$

As long as $a \neq 0$, the graph should be a conic. Describe the values of a and b that produce parabolas. What values produce ellipses? What values produce hyperbolas?

Polar Equations of Conics

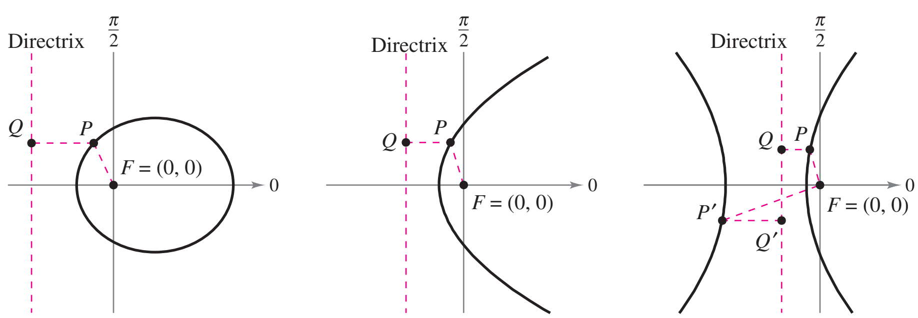
In this chapter you have seen that the rectangular equations of ellipses and hyperbolas take simple forms when the origin lies at their *centers*. As it happens, there are many important applications of conics in which it is more convenient to use one of the foci as the reference point (the origin) for the coordinate system. For example, the sun lies at a focus of Earth's orbit. Similarly, the light source of a parabolic reflector lies at its focus. In this section you will see that polar equations of conics take simple forms if one of the foci lies at the pole.

The following theorem uses the concept of *eccentricity*, as defined in Section 10.1, to classify the three basic types of conics. A proof of this theorem is given in Appendix A.

THEOREM 10.16 Classification of Conics by Eccentricity

Let F be a fixed point (*focus*) and D be a fixed line (*directrix*) in the plane. Let P be another point in the plane and let e (*eccentricity*) be the ratio of the distance between P and F to the distance between P and D. The collection of all points P with a given eccentricity is a conic.

1. The conic is an ellipse if $0 < e < 1$.
2. The conic is a parabola if $e = 1$.
3. The conic is a hyperbola if $e > 1$.



Ellipse: $0 < e < 1$
$\frac{PF}{PQ} < 1$

Parabola: $e = 1$
$PF = PQ$

Hyperbola: $e > 1$
$\frac{PF}{PQ} = \frac{P'F}{P'Q'} > 1$

Figure 10.58

In Figure 10.58, note that for each type of conic the pole corresponds to the fixed point (focus) given in the definition. The benefit of this location can be seen in the proof of the following theorem.

THEOREM 10.17 Polar Equations of Conics

The graph of a polar equation of the form

$$r = \frac{ed}{1 \pm e\cos\theta} \quad \text{or} \quad r = \frac{ed}{1 \pm e\sin\theta}$$

is a conic, where $e > 0$ is the eccentricity and $|d|$ is the distance between the focus at the pole and its corresponding directrix.

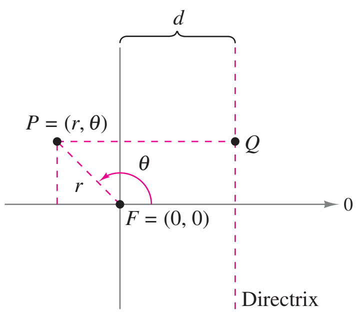


Figure 10.59

Proof The following is a proof for $r = ed/(1 + e\cos\theta)$ with $d > 0$. In Figure 10.59, consider a vertical directrix d units to the right of the focus $F = (0, 0)$. If $P = (r, \theta)$ is a point on the graph of $r = ed/(1 + e\cos\theta)$, the distance between P and the directrix can be shown to be

$$PQ = |d - x| = |d - r\cos\theta| = \left|\frac{r(1 + e\cos\theta)}{e} - r\cos\theta\right| = \left|\frac{r}{e}\right|.$$

Because the distance between P and the pole is simply $PF = |r|$, the ratio of PF to PQ is $PF/PQ = |r|/|r/e| = |e| = e$ and, by Theorem 10.16, the graph of the equation must be a conic. The proofs of the other cases are similar.

The four types of equations indicated in Theorem 10.17 can be classified as follows, where $d > 0$.

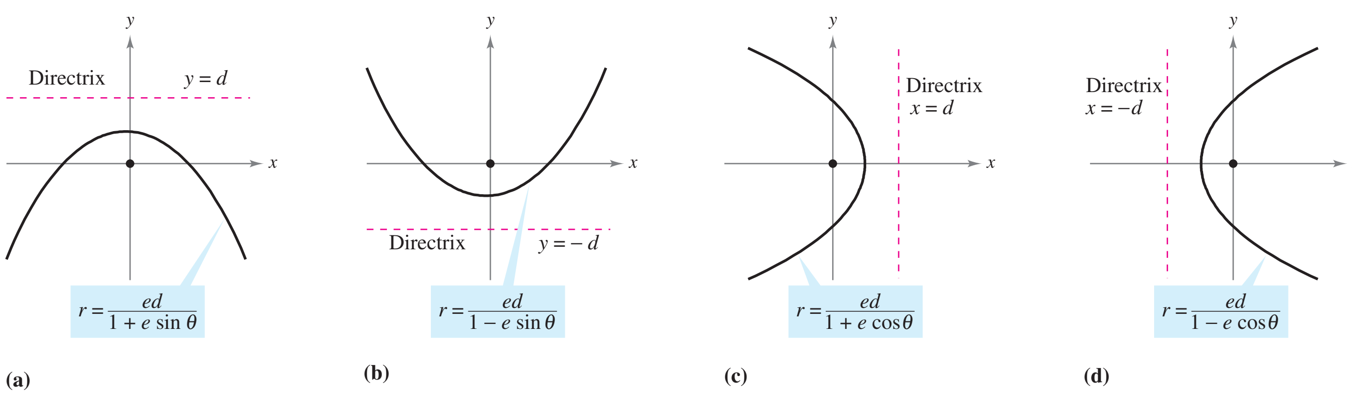
a. Horizontal directrix above the pole: $r = \dfrac{ed}{1 + e\sin\theta}$

b. Horizontal directrix below the pole: $r = \dfrac{ed}{1 - e\sin\theta}$

c. Vertical directrix to the right of the pole: $r = \dfrac{ed}{1 + e\cos\theta}$

d. Vertical directrix to the left of the pole: $r = \dfrac{ed}{1 - e\cos\theta}$

Figure 10.60 illustrates these four possibilities for a parabola.



The four types of polar equations for a parabola
Figure 10.60

EXAMPLE 1 Determining a Conic from Its Equation

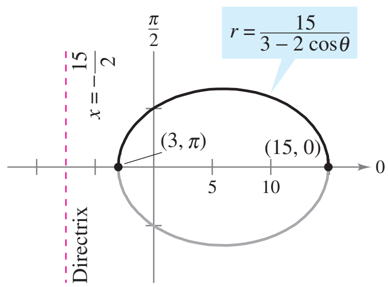
Sketch the graph of the conic given by $r = \dfrac{15}{3 - 2\cos\theta}$.

Solution To determine the type of conic, rewrite the equation as

$$r = \frac{15}{3 - 2\cos\theta} \qquad \text{Write original equation.}$$

$$= \frac{5}{1 - (2/3)\cos\theta}. \qquad \text{Divide numerator and denominator by 3.}$$

So, the graph is an ellipse with $e = \frac{2}{3}$. You can sketch the upper half of the ellipse by plotting points from $\theta = 0$ to $\theta = \pi$, as shown in Figure 10.61. Then, using symmetry with respect to the polar axis, you can sketch the lower half.



The graph of the conic is an ellipse with $e = \frac{2}{3}$.
Figure 10.61

Editable Graph

Try It Exploration A

For the ellipse in Figure 10.61, the major axis is horizontal and the vertices lie at $(15, 0)$ and $(3, \pi)$. So, the length of the *major* axis is $2a = 18$. To find the length of the minor axis, you can use the equations $e = c/a$ and $b^2 = a^2 - c^2$ to conclude

$$b^2 = a^2 - c^2 = a^2 - (ea)^2 = a^2(1 - e^2). \qquad \text{Ellipse}$$

Because $e = \frac{2}{3}$, you have

$$b^2 = 9^2\left[1 - \left(\tfrac{2}{3}\right)^2\right] = 45$$

which implies that $b = \sqrt{45} = 3\sqrt{5}$. So, the length of the minor axis is $2b = 6\sqrt{5}$. A similar analysis for hyperbolas yields

$$b^2 = c^2 - a^2 = (ea)^2 - a^2 = a^2(e^2 - 1). \qquad \text{Hyperbola}$$

EXAMPLE 2 Sketching a Conic from Its Polar Equation

Sketch the graph of the polar equation $r = \dfrac{32}{3 + 5\sin\theta}$.

Solution Dividing the numerator and denominator by 3 produces

$$r = \frac{32/3}{1 + (5/3)\sin\theta}.$$

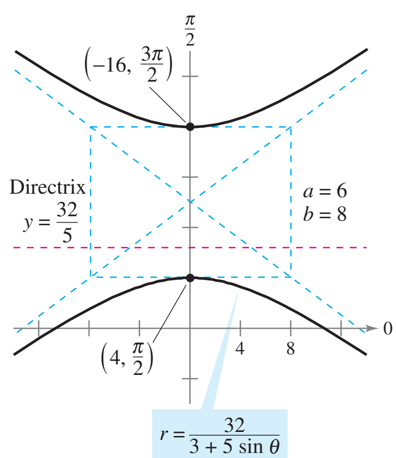
Because $e = \frac{5}{3} > 1$, the graph is a hyperbola. Because $d = \frac{32}{5}$, the directrix is the line $y = \frac{32}{5}$. The transverse axis of the hyperbola lies on the line $\theta = \pi/2$, and the vertices occur at

$$(r, \theta) = \left(4, \frac{\pi}{2}\right) \quad \text{and} \quad (r, \theta) = \left(-16, \frac{3\pi}{2}\right).$$

Because the length of the transverse axis is 12, you can see that $a = 6$. To find b, write

$$b^2 = a^2(e^2 - 1) = 6^2\left[\left(\frac{5}{3}\right)^2 - 1\right] = 64.$$

Therefore, $b = 8$. Finally, you can use a and b to determine the asymptotes of the hyperbola and obtain the sketch shown in Figure 10.62.



The graph of the conic is a hyperbola with $e = \frac{5}{3}$.
Figure 10.62

Editable Graph

Try It Exploration A Open Exploration

JOHANNES KEPLER (1571–1630)

Kepler formulated his three laws from the extensive data recorded by Danish astronomer Tycho Brahe, and from direct observation of the orbit of Mars.

MathBio

Kepler's Laws

Kepler's Laws, named after the German astronomer Johannes Kepler, can be used to describe the orbits of the planets about the sun.

1. Each planet moves in an elliptical orbit with the sun as a focus.
2. A ray from the sun to the planet sweeps out equal areas of the ellipse in equal times.
3. The square of the period is proportional to the cube of the mean distance between the planet and the sun.*

Although Kepler derived these laws empirically, they were later validated by Newton. In fact, Newton was able to show that each law can be deduced from a set of universal laws of motion and gravitation that govern the movement of all heavenly bodies, including comets and satellites. This is shown in the next example, involving the comet named after the English mathematician and physicist Edmund Halley (1656–1742).

EXAMPLE 3 Halley's Comet

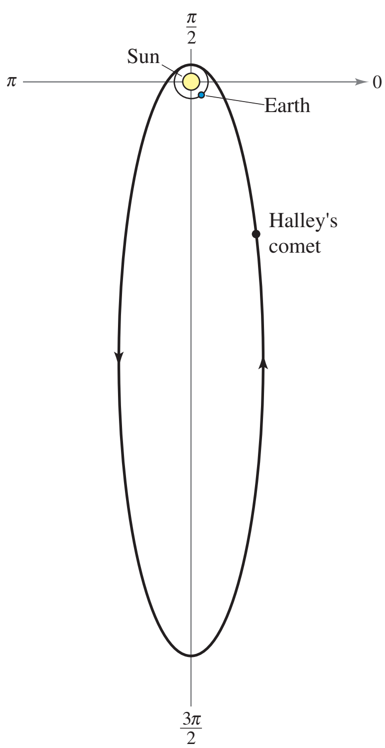


Figure 10.63

Halley's comet has an elliptical orbit with the sun at one focus and has an eccentricity of $e \approx 0.967$. The length of the major axis of the orbit is approximately 35.88 astronomical units. (An astronomical unit is defined to be the mean distance between Earth and the sun, 93 million miles.) Find a polar equation for the orbit. How close does Halley's comet come to the sun?

Solution Using a vertical axis, you can choose an equation of the form

$$r = \frac{ed}{(1 + e \sin \theta)}.$$

Because the vertices of the ellipse occur when $\theta = \pi/2$ and $\theta = 3\pi/2$, you can determine the length of the major axis to be the sum of the r-values of the vertices, as shown in Figure 10.63. That is,

$$2a = \frac{0.967d}{1 + 0.967} + \frac{0.967d}{1 - 0.967}$$

$$35.88 \approx 27.79d. \qquad 2a \approx 35.88$$

So, $d \approx 1.204$ and $ed \approx (0.967)(1.204) \approx 1.164$. Using this value in the equation produces

$$r = \frac{1.164}{1 + 0.967 \sin \theta}$$

where r is measured in astronomical units. To find the closest point to the sun (the focus), you can write $c = ea \approx (0.967)(17.94) \approx 17.35$. Because c is the distance between the focus and the center, the closest point is

$$\begin{aligned} a - c &\approx 17.94 - 17.35 \\ &\approx 0.59 \text{ AU} \\ &\approx 55{,}000{,}000 \text{ miles} \end{aligned}$$

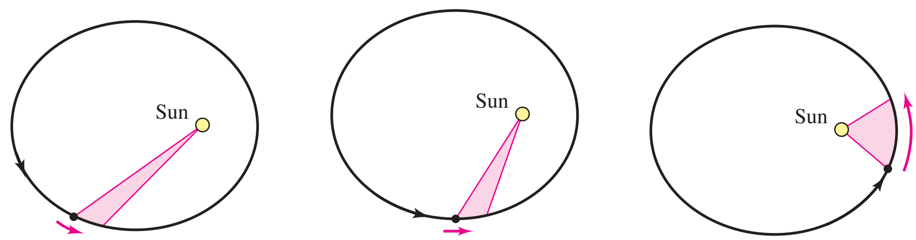
Try It Exploration A

View the video for more information about Halley's comet.

Video

* *If Earth is used as a reference with a period of 1 year and a distance of 1 astronomical unit, the proportionality constant is 1. For example, because Mars has a mean distance to the sun of $D = 1.524$ AU, its period P is given by $D^3 = P^2$. So, the period for Mars is $P = 1.88$.*

Kepler's Second Law states that as a planet moves about the sun, a ray from the sun to the planet sweeps out equal areas in equal times. This law can also be applied to comets or asteroids with elliptical orbits. For example, Figure 10.64 shows the orbit of the asteroid Apollo about the sun. Applying Kepler's Second Law to this asteroid, you know that the closer it is to the sun, the greater its velocity, because a short ray must be moving quickly to sweep out as much area as a long ray.



A ray from the sun to the asteroid sweeps out equal areas in equal times.
Figure 10.64

EXAMPLE 4 The Asteroid Apollo

The asteroid Apollo has a period of 661 Earth days, and its orbit is approximated by the ellipse

$$r = \frac{1}{1 + (5/9)\cos\theta} = \frac{9}{9 + 5\cos\theta}$$

where r is measured in astronomical units. How long does it take Apollo to move from the position given by $\theta = -\pi/2$ to $\theta = \pi/2$, as shown in Figure 10.65?

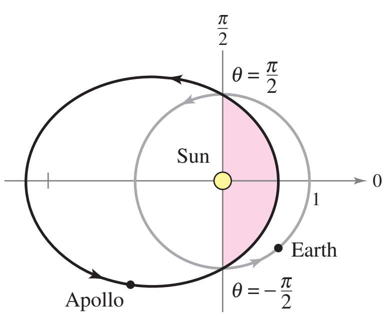


Figure 10.65

Solution Begin by finding the area swept out as θ increases from $-\pi/2$ to $\pi/2$.

$$A = \frac{1}{2}\int_{\alpha}^{\beta} r^2\, d\theta \qquad \text{Formula for area of a polar graph}$$

$$= \frac{1}{2}\int_{-\pi/2}^{\pi/2} \left(\frac{9}{9 + 5\cos\theta}\right)^2 d\theta$$

Using the substitution $u = \tan(\theta/2)$, as discussed in Section 8.6, you obtain

$$A = \frac{81}{112}\left[\frac{-5\sin\theta}{9 + 5\cos\theta} + \frac{18}{\sqrt{56}}\arctan\frac{\sqrt{56}\tan(\theta/2)}{14}\right]_{-\pi/2}^{\pi/2} \approx 0.90429.$$

Because the major axis of the ellipse has length $2a = 81/28$ and the eccentricity is $e = 5/9$, you can determine that $b = a\sqrt{1 - e^2} = 9/\sqrt{56}$. So, the area of the ellipse is

$$\text{Area of ellipse} = \pi ab = \pi\left(\frac{81}{56}\right)\left(\frac{9}{\sqrt{56}}\right) \approx 5.46507.$$

Because the time required to complete the orbit is 661 days, you can apply Kepler's Second Law to conclude that the time t required to move from the position $\theta = -\pi/2$ to $\theta = \pi/2$ is given by

$$\frac{t}{661} = \frac{\text{area of elliptical segment}}{\text{area of ellipse}} \approx \frac{0.90429}{5.46507}$$

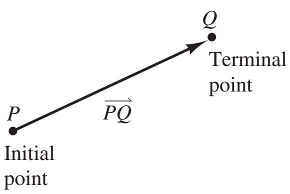
which implies that $t \approx 109$ days.

Try It Exploration A Exploration B

Section 11.1 Vectors in the Plane

- **Write the component form of a vector.**
- **Perform vector operations and interpret the results geometrically.**
- **Write a vector as a linear combination of standard unit vectors.**
- **Use vectors to solve problems involving force or velocity.**

Component Form of a Vector

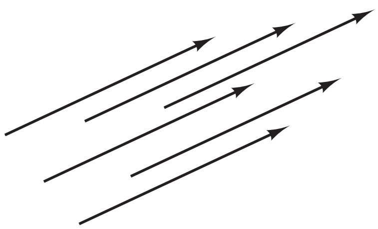


A directed line segment
Figure 11.1

Many quantities in geometry and physics, such as area, volume, temperature, mass, and time, can be characterized by a single real number scaled to appropriate units of measure. These are called **scalar quantities,** and the real number associated with each is called a **scalar.**

Other quantities, such as force, velocity, and acceleration, involve both magnitude and direction and cannot be characterized completely by a single real number. A **directed line segment** is used to represent such a quantity, as shown in Figure 11.1. The directed line segment $\overrightarrow{PQ}$ has **initial point** P and **terminal point** Q, and its **length** (or **magnitude**) is denoted by $\|\overrightarrow{PQ}\|$. Directed line segments that have the same length and direction are **equivalent,** as shown in Figure 11.2. The set of all directed line segments that are equivalent to a given directed line segment $\overrightarrow{PQ}$ is a **vector in the plane** and is denoted by $\mathbf{v} = \overrightarrow{PQ}$. In typeset material, vectors are usually denoted by lowercase, boldface letters such as **u**, **v**, and **w**. When written by hand, however, vectors are often denoted by letters with arrows above them, such as $\vec{u}$, $\vec{v}$, and $\vec{w}$.

Be sure you see that a vector in the plane can be represented by many different directed line segments—all pointing in the same direction and all of the same length.


Equivalent directed line segments
Figure 11.2

EXAMPLE 1 Vector Representation by Directed Line Segments

Let **v** be represented by the directed line segment from $(0, 0)$ to $(3, 2)$, and let **u** be represented by the directed line segment from $(1, 2)$ to $(4, 4)$. Show that **v** and **u** are equivalent.

Solution Let $P(0, 0)$ and $Q(3, 2)$ be the initial and terminal points of **v**, and let $R(1, 2)$ and $S(4, 4)$ be the initial and terminal points of **u**, as shown in Figure 11.3. You can use the Distance Formula to show that $\overrightarrow{PQ}$ and $\overrightarrow{RS}$ have the *same length*.

$$\|\overrightarrow{PQ}\| = \sqrt{(3-0)^2 + (2-0)^2} = \sqrt{13} \qquad \text{Length of } \overrightarrow{PQ}$$
$$\|\overrightarrow{RS}\| = \sqrt{(4-1)^2 + (4-2)^2} = \sqrt{13} \qquad \text{Length of } \overrightarrow{RS}$$

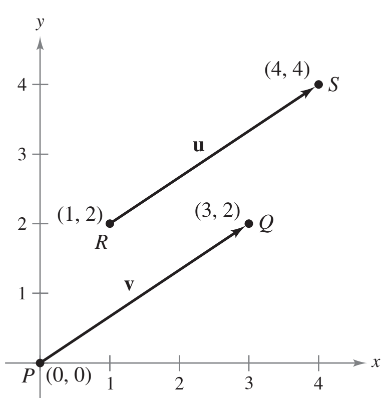
Both line segments have the *same direction*, because they both are directed toward the upper right on lines having the same slope.

$$\text{Slope of } \overrightarrow{PQ} = \frac{2-0}{3-0} = \frac{2}{3}$$

and

$$\text{Slope of } \overrightarrow{RS} = \frac{4-2}{4-1} = \frac{2}{3}$$

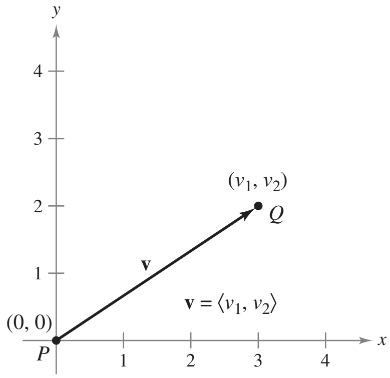
Because $\overrightarrow{PQ}$ and $\overrightarrow{RS}$ have the same length and direction, you can conclude that the two vectors are equivalent. That is, **v** and **u** are equivalent.



The vectors **u** and **v** are equivalent.
Figure 11.3

Try It

Exploration A



The standard position of a vector
Figure 11.4

The directed line segment whose initial point is the origin is often the most convenient representative of a set of equivalent directed line segments such as those shown in Figure 11.3. This representation of $\mathbf{v}$ is said to be in **standard position.** A directed line segment whose initial point is the origin can be uniquely represented by the coordinates of its terminal point $Q(v_1, v_2)$, as shown in Figure 11.4.

Definition of Component Form of a Vector in the Plane

If $\mathbf{v}$ is a vector in the plane whose initial point is the origin and whose terminal point is (v_1, v_2), then the **component form of $\mathbf{v}$** is given by

$$\mathbf{v} = \langle v_1, v_2 \rangle.$$

The coordinates v_1 and v_2 are called the **components of $\mathbf{v}$.** If both the initial point and the terminal point lie at the origin, then $\mathbf{v}$ is called the **zero vector** and is denoted by $\mathbf{0} = \langle 0, 0 \rangle$.

This definition implies that two vectors $\mathbf{u} = \langle u_1, u_2 \rangle$ and $\mathbf{v} = \langle v_1, v_2 \rangle$ are **equal** if and only if $u_1 = v_1$ and $u_2 = v_2$.

The following procedures can be used to convert directed line segments to component form or vice versa.

NOTE It is important to understand that a vector represents a *set* of directed line segments (each having the same length and direction). In practice, however, it is common not to distinguish between a vector and one of its representatives.

1. If $P(p_1, p_2)$ and $Q(q_1, q_2)$ are the initial and terminal points of a directed line segment, the component form of the vector $\mathbf{v}$ represented by $\overrightarrow{PQ}$ is $\langle v_1, v_2 \rangle = \langle q_1 - p_1, q_2 - p_2 \rangle$. Moreover, the **length** (or **magnitude**) **of v** is

$$\begin{aligned} \|\mathbf{v}\| &= \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2} \\ &= \sqrt{v_1^2 + v_2^2}. \end{aligned}$$

Length of a vector

2. If $\mathbf{v} = \langle v_1, v_2 \rangle$, $\mathbf{v}$ can be represented by the directed line segment, in standard position, from $P(0, 0)$ to $Q(v_1, v_2)$.

The length of $\mathbf{v}$ is also called the **norm of v**. If $\|\mathbf{v}\| = 1$, $\mathbf{v}$ is a **unit vector.** Moreover, $\|\mathbf{v}\| = 0$ if and only if $\mathbf{v}$ is the zero vector $\mathbf{0}$.

EXAMPLE 2 Finding the Component Form and Length of a Vector

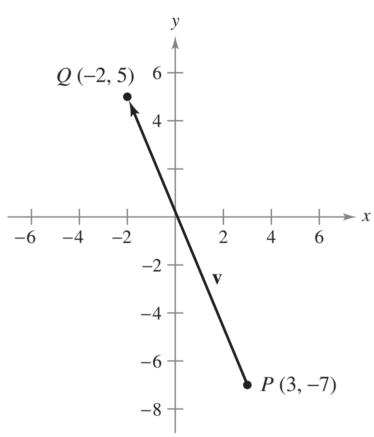
Find the component form and length of the vector $\mathbf{v}$ that has initial point $(3, -7)$ and terminal point $(-2, 5)$.

Solution Let $P(3, -7) = (p_1, p_2)$ and $Q(-2, 5) = (q_1, q_2)$. Then the components of $\mathbf{v} = \langle v_1, v_2 \rangle$ are

$$v_1 = q_1 - p_1 = -2 - 3 = -5$$
$$v_2 = q_2 - p_2 = 5 - (-7) = 12.$$

So, as shown in Figure 11.5, $\mathbf{v} = \langle -5, 12 \rangle$, and the length of $\mathbf{v}$ is

$$\begin{aligned} \|\mathbf{v}\| &= \sqrt{(-5)^2 + 12^2} \\ &= \sqrt{169} \\ &= 13. \end{aligned}$$



Component form of $\mathbf{v}$: $\mathbf{v} = \langle -5, 12 \rangle$
Figure 11.5

Try It

Exploration A

Vector Operations

Definitions of Vector Addition and Scalar Multiplication

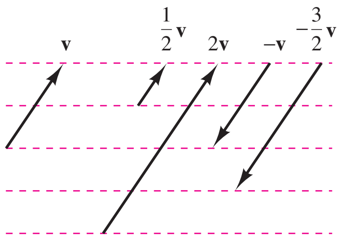
Let $\mathbf{u} = \langle u_1, u_2 \rangle$ and $\mathbf{v} = \langle v_1, v_2 \rangle$ be vectors and let c be a scalar.

1. The **vector sum** of $\mathbf{u}$ and $\mathbf{v}$ is the vector $\mathbf{u} + \mathbf{v} = \langle u_1 + v_1, u_2 + v_2 \rangle$.
2. The **scalar multiple** of c and $\mathbf{u}$ is the vector $c\mathbf{u} = \langle cu_1, cu_2 \rangle$.
3. The **negative** of $\mathbf{v}$ is the vector

$$-\mathbf{v} = (-1)\mathbf{v} = \langle -v_1, -v_2 \rangle.$$

4. The **difference** of $\mathbf{u}$ and $\mathbf{v}$ is

$$\mathbf{u} - \mathbf{v} = \mathbf{u} + (-\mathbf{v}) = \langle u_1 - v_1, u_2 - v_2 \rangle.$$



The scalar multiplication of **v**
Figure 11.6

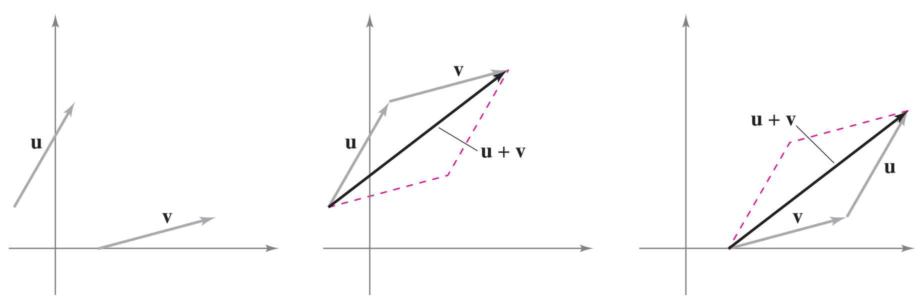
Geometrically, the scalar multiple of a vector $\mathbf{v}$ and a scalar c is the vector that is $|c|$ times as long as $\mathbf{v}$, as shown in Figure 11.6. If c is positive, $c\mathbf{v}$ has the same direction as $\mathbf{v}$. If c is negative, $c\mathbf{v}$ has the opposite direction.

The sum of two vectors can be represented geometrically by positioning the vectors (without changing their magnitudes or directions) so that the initial point of one coincides with the terminal point of the other, as shown in Figure 11.7. The vector $\mathbf{u} + \mathbf{v}$, called the **resultant vector,** is the diagonal of a parallelogram having $\mathbf{u}$ and $\mathbf{v}$ as its adjacent sides.

Isaac William Rowan Hamilton (1805–1865)

Some of the earliest work with vectors was done by the Irish mathematician William Rowan Hamilton. Hamilton spent many years developing a system of vector-like quantities called *quaternions*. Although Hamilton was convinced of the benefits of quaternions, the operations he defined did not produce good models for physical phenomena. It wasn't until the latter half of the nineteenth century that the Scottish physicist James Maxwell (1831–1879) restructured Hamilton's quaternions in a form useful for representing physical quantities such as force, velocity, and acceleration.

MathBio

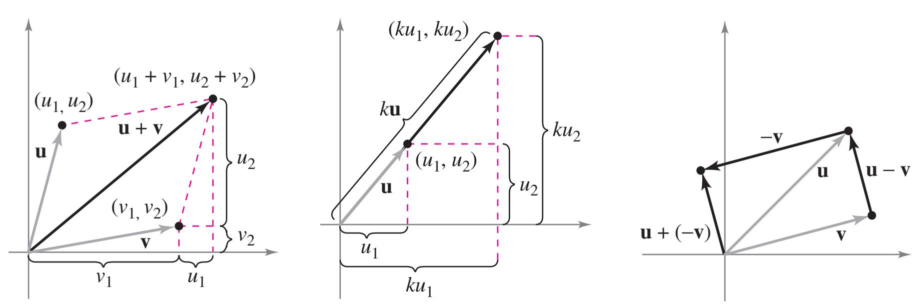


To find $\mathbf{u} + \mathbf{v}$, (1) move the initial point of **v** to the terminal point of **u**, or (2) move the initial point of **u** to the terminal point of **v**.

Figure 11.7

Animation

Figure 11.8 shows the equivalence of the geometric and algebraic definitions of vector addition and scalar multiplication, and presents (at far right) a geometric interpretation of $\mathbf{u} - \mathbf{v}$.



Vector addition Scalar multiplication Vector subtraction
Figure 11.8

EXAMPLE 3 Vector Operations

Given $\mathbf{v} = \langle -2, 5 \rangle$ and $\mathbf{w} = \langle 3, 4 \rangle$, find each of the vectors.

a. $\frac{1}{2}\mathbf{v}$ **b.** $\mathbf{w} - \mathbf{v}$ **c.** $\mathbf{v} + 2\mathbf{w}$

Solution

a. $\frac{1}{2}\mathbf{v} = \left\langle \frac{1}{2}(-2), \frac{1}{2}(5) \right\rangle = \left\langle -1, \frac{5}{2} \right\rangle$

b. $\mathbf{w} - \mathbf{v} = \langle w_1 - v_1, w_2 - v_2 \rangle = \langle 3 - (-2), 4 - 5 \rangle = \langle 5, -1 \rangle$

c. Using $2\mathbf{w} = \langle 6, 8 \rangle$, you have

$$\begin{aligned}\mathbf{v} + 2\mathbf{w} &= \langle -2, 5 \rangle + \langle 6, 8 \rangle \\ &= \langle -2 + 6, 5 + 8 \rangle \\ &= \langle 4, 13 \rangle.\end{aligned}$$

Try It | Exploration A | Exploration B

Vector addition and scalar multiplication share many properties of ordinary arithmetic, as shown in the following theorem.

THEOREM 11.1 Properties of Vector Operations

Let $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ be vectors in the plane, and let c and d be scalars.

1. $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$ — Commutative Property
2. $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$ — Associative Property
3. $\mathbf{u} + \mathbf{0} = \mathbf{u}$ — Additive Identity Property
4. $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$ — Additive Inverse Property
5. $c(d\mathbf{u}) = (cd)\mathbf{u}$
6. $(c + d)\mathbf{u} = c\mathbf{u} + d\mathbf{u}$ — Distributive Property
7. $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$ — Distributive Property
8. $1(\mathbf{u}) = \mathbf{u}$, $0(\mathbf{u}) = \mathbf{0}$

Proof The proof of the *Associative Property* of vector addition uses the Associative Property of addition of real numbers.

$$\begin{aligned}(\mathbf{u} + \mathbf{v}) + \mathbf{w} &= [\langle u_1, u_2 \rangle + \langle v_1, v_2 \rangle] + \langle w_1, w_2 \rangle \\ &= \langle u_1 + v_1, u_2 + v_2 \rangle + \langle w_1, w_2 \rangle \\ &= \langle (u_1 + v_1) + w_1, (u_2 + v_2) + w_2 \rangle \\ &= \langle u_1 + (v_1 + w_1), u_2 + (v_2 + w_2) \rangle \\ &= \langle u_1, u_2 \rangle + \langle v_1 + w_1, v_2 + w_2 \rangle = \mathbf{u} + (\mathbf{v} + \mathbf{w})\end{aligned}$$

Similarly, the proof of the *Distributive Property* of vectors depends on the Distributive Property of real numbers.

$$\begin{aligned}(c + d)\mathbf{u} &= (c + d)\langle u_1, u_2 \rangle \\ &= \langle (c + d)u_1, (c + d)u_2 \rangle \\ &= \langle cu_1 + du_1, cu_2 + du_2 \rangle \\ &= \langle cu_1, cu_2 \rangle + \langle du_1, du_2 \rangle = c\mathbf{u} + d\mathbf{u}\end{aligned}$$

The other properties can be proved in a similar manner.

Emmy Noether (1882–1935)

One person who contributed to our knowledge of axiomatic systems was the German mathematician Emmy Noether. Noether is generally recognized as the leading woman mathematician in recent history.

Any set of vectors (with an accompanying set of scalars) that satisfies the eight properties given in Theorem 11.1 is a **vector space.*** The eight properties are the *vector space axioms*. So, this theorem states that the set of vectors in the plane (with the set of real numbers) forms a vector space.

THEOREM 11.2 Length of a Scalar Multiple

Let $\mathbf{v}$ be a vector and let c be a scalar. Then

$$\|c\mathbf{v}\| = |c|\,\|\mathbf{v}\|. \qquad |c| \text{ is the absolute value of } c.$$

Proof Because $c\mathbf{v} = \langle cv_1, cv_2\rangle$, it follows that

$$\begin{aligned}
\|c\mathbf{v}\| = \|\langle cv_1, cv_2\rangle\| &= \sqrt{(cv_1)^2 + (cv_2)^2} \\
&= \sqrt{c^2v_1^2 + c^2v_2^2} \\
&= \sqrt{c^2(v_1^2 + v_2^2)} \\
&= |c|\sqrt{v_1^2 + v_2^2} \\
&= |c|\,\|\mathbf{v}\|.
\end{aligned}$$

FOR FURTHER INFORMATION For more information on Emmy Noether, see the article "Emmy Noether, Greatest Woman Mathematician" by Clark Kimberling in *The Mathematics Teacher*.

MathArticle

In many applications of vectors, it is useful to find a unit vector that has the same direction as a given vector. The following theorem gives a procedure for doing this.

THEOREM 11.3 Unit Vector in the Direction of v

If $\mathbf{v}$ is a nonzero vector in the plane, then the vector

$$\mathbf{u} = \frac{\mathbf{v}}{\|\mathbf{v}\|} = \frac{1}{\|\mathbf{v}\|}\mathbf{v}$$

has length 1 and the same direction as $\mathbf{v}$.

Proof Because $1/\|\mathbf{v}\|$ is positive and $\mathbf{u} = (1/\|\mathbf{v}\|)\mathbf{v}$, you can conclude that $\mathbf{u}$ has the same direction as $\mathbf{v}$. To see that $\|\mathbf{u}\| = 1$, note that

$$\begin{aligned}
\|\mathbf{u}\| &= \left\|\left(\frac{1}{\|\mathbf{v}\|}\right)\mathbf{v}\right\| \\
&= \left|\frac{1}{\|\mathbf{v}\|}\right|\,\|\mathbf{v}\| \\
&= \frac{1}{\|\mathbf{v}\|}\,\|\mathbf{v}\| \\
&= 1.
\end{aligned}$$

So, $\mathbf{u}$ has length 1 and the same direction as $\mathbf{v}$.

In Theorem 11.3, $\mathbf{u}$ is called a **unit vector in the direction of v.** The process of multiplying $\mathbf{v}$ by $1/\|\mathbf{v}\|$ to get a unit vector is called **normalization of v.**

* *For more information about vector spaces, see* Elementary Linear Algebra, *Fifth Edition, by Larson, Edwards, and Falvo (Boston: Houghton Mifflin Company, 2004).*

EXAMPLE 4 Finding a Unit Vector

Find a unit vector in the direction of $\mathbf{v} = \langle -2, 5 \rangle$ and verify that it has length 1.

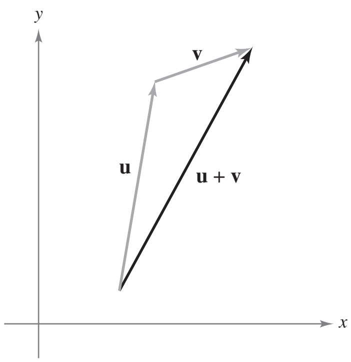
Solution From Theorem 11.3, the unit vector in the direction of $\mathbf{v}$ is

$$\frac{\mathbf{v}}{\|\mathbf{v}\|} = \frac{\langle -2, 5 \rangle}{\sqrt{(-2)^2 + (5)^2}} = \frac{1}{\sqrt{29}}\langle -2, 5 \rangle = \left\langle \frac{-2}{\sqrt{29}}, \frac{5}{\sqrt{29}} \right\rangle.$$

This vector has length 1, because

$$\sqrt{\left(\frac{-2}{\sqrt{29}}\right)^2 + \left(\frac{5}{\sqrt{29}}\right)^2} = \sqrt{\frac{4}{29} + \frac{25}{29}} = \sqrt{\frac{29}{29}} = 1.$$

Try It Exploration A



Triangle inequality
Figure 11.9

Generally, the length of the sum of two vectors is not equal to the sum of their lengths. To see this, consider the vectors $\mathbf{u}$ and $\mathbf{v}$ as shown in Figure 11.9. By considering $\mathbf{u}$ and $\mathbf{v}$ as two sides of a triangle, you can see that the length of the third side is $\|\mathbf{u} + \mathbf{v}\|$, and you have

$$\|\mathbf{u} + \mathbf{v}\| \le \|\mathbf{u}\| + \|\mathbf{v}\|.$$

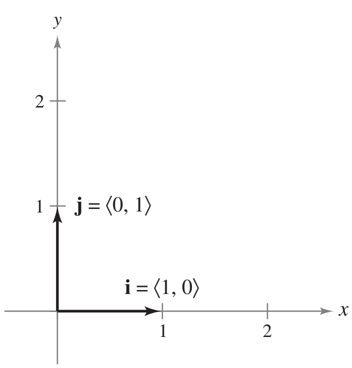
Equality occurs only if the vectors $\mathbf{u}$ and $\mathbf{v}$ have the *same direction*. This result is called the **triangle inequality** for vectors. (You are asked to prove this in Exercise 89, Section 11.3.)

Standard Unit Vectors

The unit vectors $\langle 1, 0 \rangle$ and $\langle 0, 1 \rangle$ are called the **standard unit vectors** in the plane and are denoted by

$$\mathbf{i} = \langle 1, 0 \rangle \quad \text{and} \quad \mathbf{j} = \langle 0, 1 \rangle$$

Standard unit vectors



Standard unit vectors **i** and **j**
Figure 11.10

as shown in Figure 11.10. These vectors can be used to represent any vector uniquely, as follows.

$$\mathbf{v} = \langle v_1, v_2 \rangle = \langle v_1, 0 \rangle + \langle 0, v_2 \rangle = v_1\langle 1, 0 \rangle + v_2\langle 0, 1 \rangle = v_1\mathbf{i} + v_2\mathbf{j}$$

The vector $\mathbf{v} = v_1\mathbf{i} + v_2\mathbf{j}$ is called a **linear combination** of $\mathbf{i}$ and $\mathbf{j}$. The scalars v_1 and v_2 are called the **horizontal** and **vertical components of v.**

EXAMPLE 5 Writing a Linear Combination of Unit Vectors

Let $\mathbf{u}$ be the vector with initial point $(2, -5)$ and terminal point $(-1, 3)$, and let $\mathbf{v} = 2\mathbf{i} - \mathbf{j}$. Write each vector as a linear combination of $\mathbf{i}$ and $\mathbf{j}$.

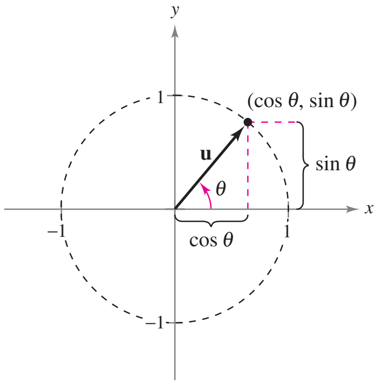
a. $\mathbf{u}$ **b.** $\mathbf{w} = 2\mathbf{u} - 3\mathbf{v}$

Solution

a. $\mathbf{u} = \langle q_1 - p_1, q_2 - p_2 \rangle$
$= \langle -1 - 2, 3 - (-5) \rangle$
$= \langle -3, 8 \rangle = -3\mathbf{i} + 8\mathbf{j}$

b. $\mathbf{w} = 2\mathbf{u} - 3\mathbf{v} = 2(-3\mathbf{i} + 8\mathbf{j}) - 3(2\mathbf{i} - \mathbf{j})$
$= -6\mathbf{i} + 16\mathbf{j} - 6\mathbf{i} + 3\mathbf{j}$
$= -12\mathbf{i} + 19\mathbf{j}$

Try It Exploration A



The angle θ from the positive x-axis to the vector **u**
Figure 11.11

If **u** is a unit vector and θ is the angle (measured counterclockwise) from the positive x-axis to **u**, then the terminal point of **u** lies on the unit circle, and you have

$$\mathbf{u} = \langle \cos\theta, \sin\theta \rangle = \cos\theta\mathbf{i} + \sin\theta\mathbf{j} \qquad \text{Unit vector}$$

as shown in Figure 11.11. Moreover, it follows that any other nonzero vector **v** making an angle θ with the positive x-axis has the same direction as **u**, and you can write

$$\mathbf{v} = \|\mathbf{v}\|\langle \cos\theta, \sin\theta \rangle = \|\mathbf{v}\|\cos\theta\mathbf{i} + \|\mathbf{v}\|\sin\theta\mathbf{j}.$$

EXAMPLE 6 Writing a Vector of Given Magnitude and Direction

The vector **v** has a magnitude of 3 and makes an angle of $30° = \pi/6$ with the positive x-axis. Write **v** as a linear combination of the unit vectors **i** and **j**.

Solution Because the angle between **v** and the positive x-axis is $\theta = \pi/6$, you can write the following.

$$\begin{aligned}\mathbf{v} &= \|\mathbf{v}\|\cos\theta\mathbf{i} + \|\mathbf{v}\|\sin\theta\mathbf{j} \\ &= 3\cos\frac{\pi}{6}\mathbf{i} + 3\sin\frac{\pi}{6}\mathbf{j} \\ &= \frac{3\sqrt{3}}{2}\mathbf{i} + \frac{3}{2}\mathbf{j}\end{aligned}$$

Try It Exploration A Exploration B

Applications of Vectors

Vectors have many applications in physics and engineering. One example is force. A vector can be used to represent force because force has both magnitude and direction. If two or more forces are acting on an object, then the **resultant force** on the object is the vector sum of the vector forces.

EXAMPLE 7 Finding the Resultant Force

Two tugboats are pushing an ocean liner, as shown in Figure 11.12. Each boat is exerting a force of 400 pounds. What is the resultant force on the ocean liner?

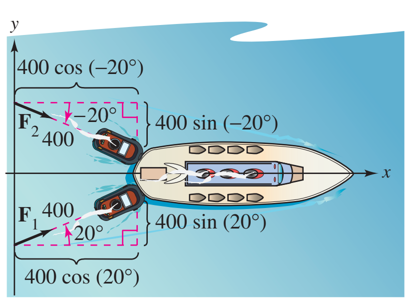
Solution Using Figure 11.12, you can represent the forces exerted by the first and second tugboats as

$$\begin{aligned}\mathbf{F}_1 &= 400\langle \cos 20°, \sin 20° \rangle \\ &= 400\cos(20°)\mathbf{i} + 400\sin(20°)\mathbf{j} \\ \mathbf{F}_2 &= 400\langle \cos(-20°), \sin(-20°) \rangle \\ &= 400\cos(20°)\mathbf{i} - 400\sin(20°)\mathbf{j}.\end{aligned}$$

The resultant force on the ocean liner is

$$\begin{aligned}\mathbf{F} &= \mathbf{F}_1 + \mathbf{F}_2 \\ &= [400\cos(20°)\mathbf{i} + 400\sin(20°)\mathbf{j}] + [400\cos(20°)\mathbf{i} - 400\sin(20°)\mathbf{j}] \\ &= 800\cos(20°)\mathbf{i} \\ &\approx 752\mathbf{i}.\end{aligned}$$

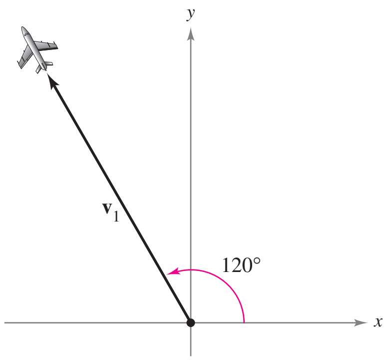
So, the resultant force on the ocean liner is approximately 752 pounds in the direction of the positive x-axis.



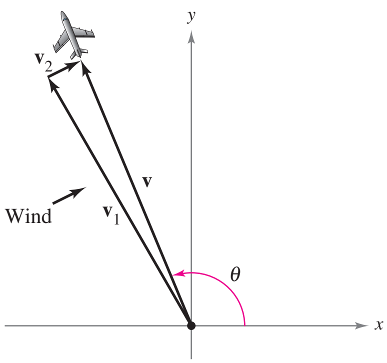
The resultant force on the ocean liner that is exerted by the two tugboats.
Figure 11.12

Try It Exploration A

In surveying and navigation, a bearing is a direction that measures the acute angle that a path or line of sight makes with a fixed north-south line. In air navigation, bearings are measured in degrees clockwise from north.



(a) Direction without wind



(b) Direction with wind
Figure 11.13

EXAMPLE 8 Finding a Velocity

An airplane is traveling at a fixed altitude with a negligible wind factor. The airplane is traveling at a speed of 500 miles per hour with a bearing of 330°, as shown in Figure 11.13(a). As the airplane reaches a certain point, it encounters wind with a velocity of 70 miles per hour in the direction N 45° E (45° east of north), as shown in Figure 11.13(b). What are the resultant speed and direction of the airplane?

Solution Using Figure 11.13(a), represent the velocity of the airplane (alone) as

$$\mathbf{v}_1 = 500\cos(120°)\mathbf{i} + 500\sin(120°)\mathbf{j}.$$

The velocity of the wind is represented by the vector

$$\mathbf{v}_2 = 70\cos(45°)\mathbf{i} + 70\sin(45°)\mathbf{j}.$$

The resultant velocity of the airplane (in the wind) is

$$\begin{aligned}\mathbf{v} = \mathbf{v}_1 + \mathbf{v}_2 &= 500\cos(120°)\mathbf{i} + 500\sin(120°)\mathbf{j} + 70\cos(45°)\mathbf{i} + 70\sin(45°)\mathbf{j}\\ &\approx -200.5\mathbf{i} + 482.5\mathbf{j}.\end{aligned}$$

To find the resultant speed and direction, write $\mathbf{v} = \|\mathbf{v}\|(\cos\theta\,\mathbf{i} + \sin\theta\,\mathbf{j})$. Because $\|\mathbf{v}\| \approx \sqrt{(-200.5)^2 + (482.5)^2} \approx 522.5$, you can write

$$\mathbf{v} \approx 522.5\left(\frac{-200.5}{522.5}\mathbf{i} + \frac{482.5}{522.5}\mathbf{j}\right) \approx 522.5,[\cos(112.6°)\mathbf{i} + \sin(112.6°)\mathbf{j}].$$

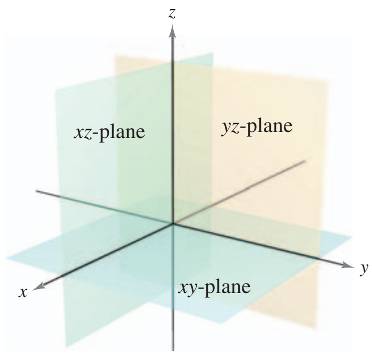
The new speed of the airplane, as altered by the wind, is approximately 522.5 miles per hour in a path that makes an angle of 112.6° with the positive x-axis.

Try It   Open Exploration

Section 11.2 Space Coordinates and Vectors in Space

- **Understand the three-dimensional rectangular coordinate system.**
- **Analyze vectors in space.**
- **Use three-dimensional vectors to solve real-life problems.**

Coordinates in Space



The three-dimensional coordinate system
Figure 11.14

Rotatable Graph

Up to this point in the text, you have been primarily concerned with the two-dimensional coordinate system. Much of the remaining part of your study of calculus will involve the three-dimensional coordinate system.

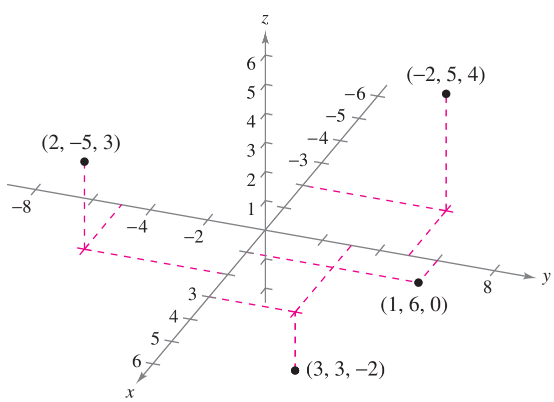
Before extending the concept of a vector to three dimensions, you must be able to identify points in the **three-dimensional coordinate system.** You can construct this system by passing a z-axis perpendicular to both the x- and y-axes at the origin. Figure 11.14 shows the positive portion of each coordinate axis. Taken as pairs, the axes determine three **coordinate planes:** the ***xy*-plane,** the ***xz*-plane,** and the ***yz*-plane.** These three coordinate planes separate three-space into eight **octants.** The first octant is the one for which all three coordinates are positive. In this three-dimensional system, a point P in space is determined by an ordered triple (x, y, z) where x, y, and z are as follows.

$x =$ directed distance from yz-plane to P

$y =$ directed distance from xz-plane to P

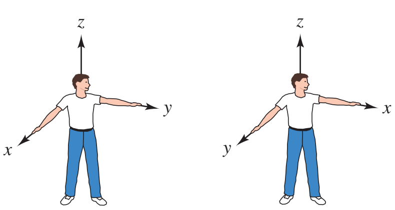
$z =$ directed distance from xy-plane to P

Several points are shown in Figure 11.15.



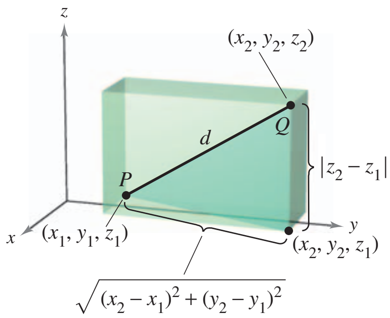
Points in the three-dimensional coordinate system are represented by ordered triples.
Figure 11.15

Rotatable Graph



Right-handed system Left-handed system
Figure 11.16

A three-dimensional coordinate system can have either a **left-handed** or a **right-handed** orientation. To determine the orientation of a system, imagine that you are standing at the origin, with your arms pointing in the direction of the positive x- and y-axes, and with the z-axis pointing up, as shown in Figure 11.16. The system is right-handed or left-handed depending on which hand points along the x-axis. In this text, you will work exclusively with the right-handed system.



The distance between two points in space
Figure 11.17

Rotatable Graph

Many of the formulas established for the two-dimensional coordinate system can be extended to three dimensions. For example, to find the distance between two points in space, you can use the Pythagorean Theorem twice, as shown in Figure 11.17. By doing this, you will obtain the formula for the distance between the points (x_1, y_1, z_1) and (x_2, y_2, z_2).

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}$$ Distance Formula

EXAMPLE 1 Finding the Distance Between Two Points in Space

The distance between the points $(2, -1, 3)$ and $(1, 0, -2)$ is

$$\begin{aligned} d &= \sqrt{(1 - 2)^2 + (0 + 1)^2 + (-2 - 3)^2} && \text{Distance Formula} \\ &= \sqrt{1 + 1 + 25} \\ &= \sqrt{27} \\ &= 3\sqrt{3}. \end{aligned}$$

Try It Exploration A Exploration B

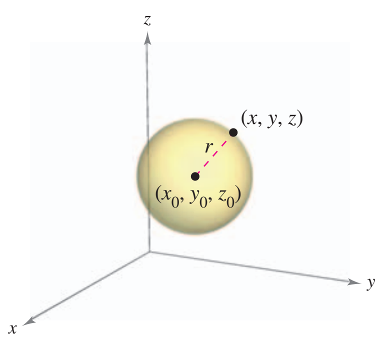


Figure 11.18

Rotatable Graph

A **sphere** with center at (x_0, y_0, z_0) and radius r is defined to be the set of all points (x, y, z) such that the distance between (x, y, z) and (x_0, y_0, z_0) is r. You can use the Distance Formula to find the **standard equation of a sphere** of radius r, centered at (x_0, y_0, z_0). If (x, y, z) is an arbitrary point on the sphere, the equation of the sphere is

$$(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2 = r^2$$ Equation of sphere

as shown in Figure 11.18. Moreover, the midpoint of the line segment joining the points (x_1, y_1, z_1) and (x_2, y_2, z_2) has coordinates

$$\left(\frac{x_1 + x_2}{2}, \frac{y_1 + y_2}{2}, \frac{z_1 + z_2}{2}\right).$$ Midpoint Rule

EXAMPLE 2 Finding the Equation of a Sphere

Find the standard equation of the sphere that has the points $(5, -2, 3)$ and $(0, 4, -3)$ as endpoints of a diameter.

Solution By the Midpoint Rule, the center of the sphere is

$$\left(\frac{5 + 0}{2}, \frac{-2 + 4}{2}, \frac{3 - 3}{2}\right) = \left(\frac{5}{2}, 1, 0\right).$$ Midpoint Rule

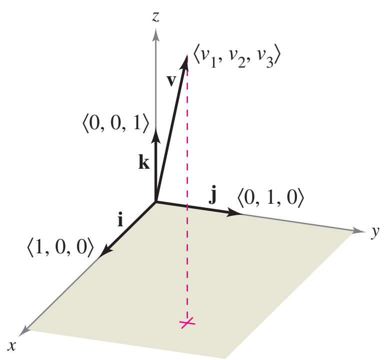
By the Distance Formula, the radius is

$$r = \sqrt{\left(0 - \frac{5}{2}\right)^2 + (4 - 1)^2 + (-3 - 0)^2} = \sqrt{\frac{97}{4}} = \frac{\sqrt{97}}{2}.$$

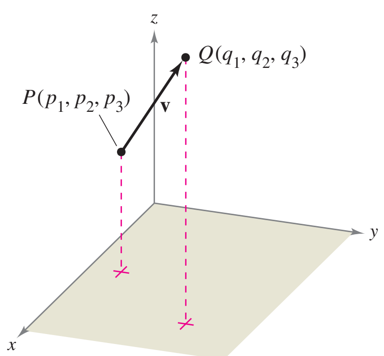
Therefore, the standard equation of the sphere is

$$\left(x - \frac{5}{2}\right)^2 + (y - 1)^2 + z^2 = \frac{97}{4}.$$ Equation of sphere

Try It Exploration A



The standard unit vectors in space
Figure 11.19



$\mathbf{v} = \langle q_1 - p_1, q_2 - p_2, q_3 - p_3\rangle$

Figure 11.20

Vectors in Space

In space, vectors are denoted by ordered triples $\mathbf{v} = \langle v_1, v_2, v_3\rangle$. The **zero vector** is denoted by $\mathbf{0} = \langle 0, 0, 0\rangle$. Using the unit vectors $\mathbf{i} = \langle 1, 0, 0\rangle$, $\mathbf{j} = \langle 0, 1, 0\rangle$, and $\mathbf{k} = \langle 0, 0, 1\rangle$ in the direction of the positive z-axis, the **standard unit vector notation** for $\mathbf{v}$ is

$$\mathbf{v} = v_1\mathbf{i} + v_2\mathbf{j} + v_3\mathbf{k}$$

as shown in Figure 11.19. If $\mathbf{v}$ is represented by the directed line segment from $P(p_1, p_2, p_3)$ to $Q(q_1, q_2, q_3)$, as shown in Figure 11.20, the component form of $\mathbf{v}$ is given by subtracting the coordinates of the initial point from the coordinates of the terminal point, as follows.

$$\mathbf{v} = \langle v_1, v_2, v_3\rangle = \langle q_1 - p_1, q_2 - p_2, q_3 - p_3\rangle$$

Vectors in Space

Let $\mathbf{u} = \langle u_1, u_2, u_3\rangle$ and $\mathbf{v} = \langle v_1, v_2, v_3\rangle$ be vectors in space and let c be a scalar.

1. *Equality of Vectors:* $\mathbf{u} = \mathbf{v}$ if and only if $u_1 = v_1$, $u_2 = v_2$, and $u_3 = v_3$.
2. *Component Form:* If $\mathbf{v}$ is represented by the directed line segment from $P(p_1, p_2, p_3)$ to $Q(q_1, q_2, q_3)$, then
$$\mathbf{v} = \langle v_1, v_2, v_3\rangle = \langle q_1 - p_1, q_2 - p_2, q_3 - p_3\rangle.$$
3. *Length:* $\|\mathbf{v}\| = \sqrt{v_1^2 + v_2^2 + v_3^2}$
4. *Unit Vector in the Direction of* $\mathbf{v}$*:* $\dfrac{\mathbf{v}}{\|\mathbf{v}\|} = \left(\dfrac{1}{\|\mathbf{v}\|}\right)\langle v_1, v_2, v_3\rangle, \quad \mathbf{v} \neq \mathbf{0}$
5. *Vector Addition:* $\mathbf{v} + \mathbf{u} = \langle v_1 + u_1, v_2 + u_2, v_3 + u_3\rangle$
6. *Scalar Multiplication:* $c\mathbf{v} = \langle cv_1, cv_2, cv_3\rangle$

NOTE The properties of vector addition and scalar multiplication given in Theorem 11.1 are also valid for vectors in space.

EXAMPLE 3 Finding the Component Form of a Vector in Space

Find the component form and magnitude of the vector $\mathbf{v}$ having initial point $(-2, 3, 1)$ and terminal point $(0, -4, 4)$. Then find a unit vector in the direction of $\mathbf{v}$.

Solution The component form of $\mathbf{v}$ is

$$\begin{aligned}\mathbf{v} = \langle q_1 - p_1, q_2 - p_2, q_3 - p_3\rangle &= \langle 0 - (-2), -4 - 3, 4 - 1\rangle \\ &= \langle 2, -7, 3\rangle\end{aligned}$$

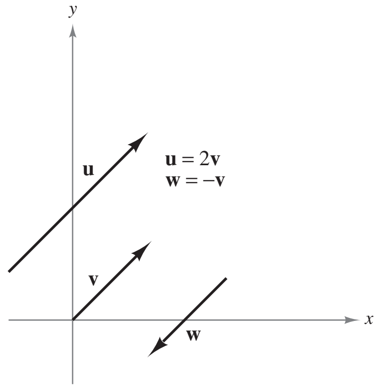
which implies that its magnitude is

$$\|\mathbf{v}\| = \sqrt{2^2 + (-7)^2 + 3^2} = \sqrt{62}.$$

The unit vector in the direction of $\mathbf{v}$ is

$$\mathbf{u} = \frac{\mathbf{v}}{\|\mathbf{v}\|} = \frac{1}{\sqrt{62}}\langle 2, -7, 3\rangle.$$

Try It | Exploration A | Exploration B | Open Exploration



Parallel vectors
Figure 11.21

Recall from the definition of scalar multiplication that positive scalar multiples of a nonzero vector **v** have the same direction as **v**, whereas negative multiples have the direction opposite of **v**. In general, two nonzero vectors **u** and **v** are **parallel** if there is some scalar c such that $\mathbf{u} = c\mathbf{v}$.

Definition of Parallel Vectors

Two nonzero vectors **u** and **v** are **parallel** if there is some scalar c such that $\mathbf{u} = c\mathbf{v}$.

For example, in Figure 11.21, the vectors **u**, **v**, and **w** are parallel because $\mathbf{u} = 2\mathbf{v}$ and $\mathbf{w} = -\mathbf{v}$.

EXAMPLE 4 Parallel Vectors

Vector **w** has initial point $(2, -1, 3)$ and terminal point $(-4, 7, 5)$. Which of the following vectors is parallel to **w**?

a. $\mathbf{u} = \langle 3, -4, -1\rangle$

b. $\mathbf{v} = \langle 12, -16, 4\rangle$

Solution Begin by writing **w** in component form.

$$\mathbf{w} = \langle -4 - 2, 7 - (-1), 5 - 3\rangle = \langle -6, 8, 2\rangle$$

a. Because $\mathbf{u} = \langle 3, -4, -1\rangle = -\frac{1}{2}\langle -6, 8, 2\rangle = -\frac{1}{2}\mathbf{w}$, you can conclude that **u** is parallel to **w**.

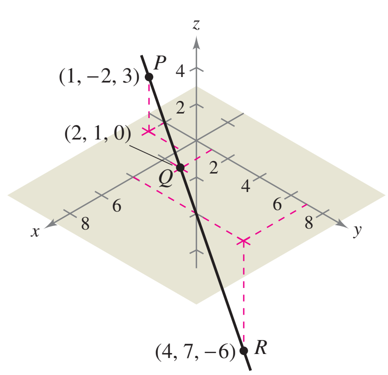
b. In this case, you want to find a scalar c such that

$$\langle 12, -16, 4\rangle = c\langle -6, 8, 2\rangle.$$

$$\begin{aligned} 12 &= -6c \rightarrow c = -2 \\ -16 &= 8c \rightarrow c = -2 \\ 4 &= 2c \rightarrow c = 2 \end{aligned}$$

Because there is no c for which the equation has a solution, the vectors are not parallel.

Try It   Exploration A



The points P, Q, and R lie on the same line.
Figure 11.22

EXAMPLE 5 Using Vectors to Determine Collinear Points

Determine whether the points $P(1, -2, 3)$, $Q(2, 1, 0)$, and $R(4, 7, -6)$ are collinear.

Solution The component forms of $\overrightarrow{PQ}$ and $\overrightarrow{PR}$ are

$$\overrightarrow{PQ} = \langle 2 - 1, 1 - (-2), 0 - 3\rangle = \langle 1, 3, -3\rangle$$

and

$$\overrightarrow{PR} = \langle 4 - 1, 7 - (-2), -6 - 3\rangle = \langle 3, 9, -9\rangle.$$

These two vectors have a common initial point. So, P, Q, and R lie on the same line if and only if $\overrightarrow{PQ}$ and $\overrightarrow{PR}$ are parallel—which they are because $\overrightarrow{PR} = 3\,\overrightarrow{PQ}$, as shown in Figure 11.22.

Try It   Exploration A

EXAMPLE 6 Standard Unit Vector Notation

a. Write the vector $\mathbf{v} = 4\mathbf{i} - 5\mathbf{k}$ in component form.

b. Find the terminal point of the vector $\mathbf{v} = 7\mathbf{i} - \mathbf{j} + 3\mathbf{k}$, given that the initial point is $P(-2, 3, 5)$.

Solution

a. Because $\mathbf{j}$ is missing, its component is 0 and

$$\mathbf{v} = 4\mathbf{i} - 5\mathbf{k} = \langle 4, 0, -5\rangle.$$

b. You need to find $Q(q_1, q_2, q_3)$ such that $\mathbf{v} = \overrightarrow{PQ} = 7\mathbf{i} - \mathbf{j} + 3\mathbf{k}$. This implies that $q_1 - (-2) = 7$, $q_2 - 3 = -1$, and $q_3 - 5 = 3$. The solution of these three equations is $q_1 = 5$, $q_2 = 2$, and $q_3 = 8$. Therefore, Q is $(5, 2, 8)$.

Try It Exploration A

Application

EXAMPLE 7 Measuring Force

A television camera weighing 120 pounds is supported by a tripod, as shown in Figure 11.23. Represent the force exerted on each leg of the tripod as a vector.

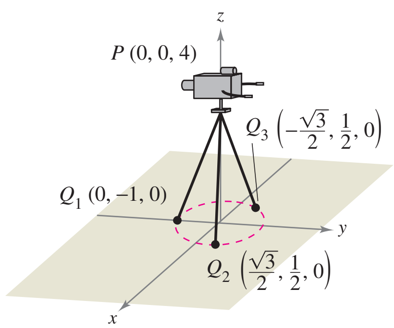


Figure 11.23

Solution Let the vectors $\mathbf{F}_1$, $\mathbf{F}_2$, and $\mathbf{F}_3$ represent the forces exerted on the three legs. From Figure 11.23, you can determine the directions of $\mathbf{F}_1$, $\mathbf{F}_2$, and $\mathbf{F}_3$ to be as follows.

$$\overrightarrow{PQ}_1 = \langle 0 - 0, -1 - 0, 0 - 4\rangle = \langle 0, -1, -4\rangle$$
$$\overrightarrow{PQ}_2 = \left\langle \frac{\sqrt{3}}{2} - 0, \frac{1}{2} - 0, 0 - 4\right\rangle = \left\langle \frac{\sqrt{3}}{2}, \frac{1}{2}, -4\right\rangle$$
$$\overrightarrow{PQ}_3 = \left\langle -\frac{\sqrt{3}}{2} - 0, \frac{1}{2} - 0, 0 - 4\right\rangle = \left\langle -\frac{\sqrt{3}}{2}, \frac{1}{2}, -4\right\rangle$$

Because each leg has the same length, and the total force is distributed equally among the three legs, you know that $\|\mathbf{F}_1\| = \|\mathbf{F}_2\| = \|\mathbf{F}_3\|$. So, there exists a constant c such that

$$\mathbf{F}_1 = c\langle 0, -1, -4\rangle, \quad \mathbf{F}_2 = c\left\langle \frac{\sqrt{3}}{2}, \frac{1}{2}, -4\right\rangle, \quad \text{and} \quad \mathbf{F}_3 = c\left\langle -\frac{\sqrt{3}}{2}, \frac{1}{2}, -4\right\rangle.$$

Let the total force exerted by the object be given by $\mathbf{F} = -120\mathbf{k}$. Then, using the fact that

$$\mathbf{F} = \mathbf{F}_1 + \mathbf{F}_2 + \mathbf{F}_3$$

you can conclude that $\mathbf{F}_1$, $\mathbf{F}_2$, and $\mathbf{F}_3$ all have a vertical component of -40. This implies that $c(-4) = -40$ and $c = 10$. Therefore, the forces exerted on the legs can be represented by

$$\mathbf{F}_1 = \langle 0, -10, -40\rangle$$
$$\mathbf{F}_2 = \langle 5\sqrt{3}, 5, -40\rangle$$
$$\mathbf{F}_3 = \langle -5\sqrt{3}, 5, -40\rangle.$$

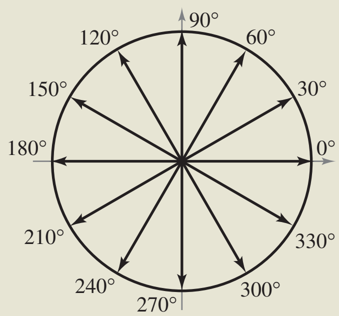
Try It Exploration A

Section 11.3 The Dot Product of Two Vectors

- **Use properties of the dot product of two vectors.**
- **Find the angle between two vectors using the dot product.**
- **Find the direction cosines of a vector in space.**
- **Find the projection of a vector onto another vector.**
- **Use vectors to find the work done by a constant force.**

EXPLORATION

Interpreting a Dot Product Several vectors are shown below on the unit circle. Find the dot products of several pairs of vectors. Then find the angle between each pair that you used. Make a conjecture about the relationship between the dot product of two vectors and the angle between the vectors.



The Dot Product

So far you have studied two operations with vectors—vector addition and multiplication by a scalar—each of which yields another vector. In this section you will study a third vector operation, called the **dot product.** This product yields a scalar, rather than a vector.

Definition of Dot Product

The **dot product** of $\mathbf{u} = \langle u_1, u_2\rangle$ and $\mathbf{v} = \langle v_1, v_2\rangle$ is

$$\mathbf{u} \cdot \mathbf{v} = u_1v_1 + u_2v_2.$$

The **dot product** of $\mathbf{u} = \langle u_1, u_2, u_3\rangle$ and $\mathbf{v} = \langle v_1, v_2, v_3\rangle$ is

$$\mathbf{u} \cdot \mathbf{v} = u_1v_1 + u_2v_2 + u_3v_3.$$

NOTE Because the dot product of two vectors yields a scalar, it is also called the **inner product** (or **scalar product**) of the two vectors.

THEOREM 11.4 Properties of the Dot Product

Let $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ be vectors in the plane or in space and let c be a scalar.

1. $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$ Commutative Property
2. $\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w}$ Distributive Property
3. $c(\mathbf{u} \cdot \mathbf{v}) = c\mathbf{u} \cdot \mathbf{v} = \mathbf{u} \cdot c\mathbf{v}$
4. $\mathbf{0} \cdot \mathbf{v} = 0$
5. $\mathbf{v} \cdot \mathbf{v} = \|\mathbf{v}\|^2$

Proof To prove the first property, let $\mathbf{u} = \langle u_1, u_2, u_3\rangle$ and $\mathbf{v} = \langle v_1, v_2, v_3\rangle$. Then

$$\begin{aligned}\mathbf{u} \cdot \mathbf{v} &= u_1v_1 + u_2v_2 + u_3v_3\\ &= v_1u_1 + v_2u_2 + v_3u_3\\ &= \mathbf{v} \cdot \mathbf{u}.\end{aligned}$$

For the fifth property, let $\mathbf{v} = \langle v_1, v_2, v_3\rangle$. Then

$$\begin{aligned}\mathbf{v} \cdot \mathbf{v} &= v_1^2 + v_2^2 + v_3^2\\ &= \left(\sqrt{v_1^2 + v_2^2 + v_3^2}\right)^2\\ &= \|\mathbf{v}\|^2.\end{aligned}$$

Proofs of the other properties are left to you.

EXAMPLE 1 Finding Dot Products

Given $\mathbf{u} = \langle 2, -2 \rangle$, $\mathbf{v} = \langle 5, 8 \rangle$, and $\mathbf{w} = \langle -4, 3 \rangle$, find each of the following.

a. $\mathbf{u} \cdot \mathbf{v}$ **b.** $(\mathbf{u} \cdot \mathbf{v})\mathbf{w}$

c. $\mathbf{u} \cdot (2\mathbf{v})$ **d.** $\|\mathbf{w}\|^2$

Solution

a. $\mathbf{u} \cdot \mathbf{v} = \langle 2, -2 \rangle \cdot \langle 5, 8 \rangle = 2(5) + (-2)(8) = -6$

b. $(\mathbf{u} \cdot \mathbf{v})\mathbf{w} = -6\langle -4, 3 \rangle = \langle 24, -18 \rangle$

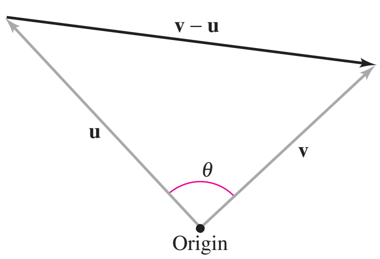
c. $\mathbf{u} \cdot (2\mathbf{v}) = 2(\mathbf{u} \cdot \mathbf{v}) = 2(-6) = -12$ Theorem 11.4

d.
$$\begin{aligned}
\|\mathbf{w}\|^2 &= \mathbf{w} \cdot \mathbf{w} && \text{Theorem 11.4}\\
&= \langle -4, 3 \rangle \cdot \langle -4, 3 \rangle && \text{Substitute } \langle -4, 3 \rangle \text{ for } \mathbf{w}.\\
&= (-4)(-4) + (3)(3) && \text{Definition of dot product}\\
&= 25 && \text{Simplify.}
\end{aligned}$$

Notice that the result of part (b) is a *vector* quantity, whereas the results of the other three parts are *scalar* quantities.

Try It Exploration A Exploration B

Angle Between Two Vectors



The angle between two vectors
Figure 11.24

The **angle between two nonzero vectors** is the angle θ, $0 \le \theta \le \pi$, between their respective standard position vectors, as shown in Figure 11.24. The next theorem shows how to find this angle using the dot product. (Note that the angle between the zero vector and another vector is not defined here.)

THEOREM 11.5 Angle Between Two Vectors

If θ is the angle between two nonzero vectors $\mathbf{u}$ and $\mathbf{v}$, then

$$\cos\theta = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\|\,\|\mathbf{v}\|}.$$

Proof Consider the triangle determined by vectors $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{v} - \mathbf{u}$, as shown in Figure 11.24. By the Law of Cosines, you can write

$$\|\mathbf{v} - \mathbf{u}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - 2\|\mathbf{u}\|\,\|\mathbf{v}\|\cos\theta.$$

Using the properties of the dot product, the left side can be rewritten as

$$\begin{aligned}
\|\mathbf{v} - \mathbf{u}\|^2 &= (\mathbf{v} - \mathbf{u}) \cdot (\mathbf{v} - \mathbf{u})\\
&= (\mathbf{v} - \mathbf{u}) \cdot \mathbf{v} - (\mathbf{v} - \mathbf{u}) \cdot \mathbf{u}\\
&= \mathbf{v} \cdot \mathbf{v} - \mathbf{u} \cdot \mathbf{v} - \mathbf{v} \cdot \mathbf{u} + \mathbf{u} \cdot \mathbf{u}\\
&= \|\mathbf{v}\|^2 - 2\mathbf{u} \cdot \mathbf{v} + \|\mathbf{u}\|^2
\end{aligned}$$

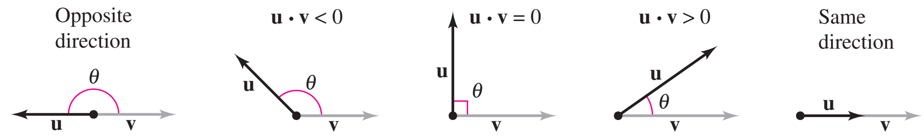
and substitution back into the Law of Cosines yields

$$\begin{aligned}
\|\mathbf{v}\|^2 - 2\mathbf{u} \cdot \mathbf{v} + \|\mathbf{u}\|^2 &= \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - 2\|\mathbf{u}\|\,\|\mathbf{v}\|\cos\theta\\
-2\mathbf{u} \cdot \mathbf{v} &= -2\|\mathbf{u}\|\,\|\mathbf{v}\|\cos\theta\\
\cos\theta &= \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\|\,\|\mathbf{v}\|}.
\end{aligned}$$

If the angle between two vectors is known, rewriting Theorem 11.5 in the form

$$\mathbf{u} \cdot \mathbf{v} = \|\mathbf{u}\|\,\|\mathbf{v}\| \cos \theta \qquad \text{Alternative form of dot product}$$

produces an alternative way to calculate the dot product. From this form, you can see that because $\|\mathbf{u}\|$ and $\|\mathbf{v}\|$ are always positive, $\mathbf{u} \cdot \mathbf{v}$ and $\cos \theta$ will always have the same sign. Figure 11.25 shows the possible orientations of two vectors.



$\theta = \pi$	$\pi/2 < \theta < \pi$	$\theta = \pi/2$	$0 < \theta < \pi/2$	$\theta = 0$
$\cos \theta = -1$	$-1 < \cos \theta < 0$	$\cos \theta = 0$	$0 < \cos \theta < 1$	$\cos \theta = 1$

Figure 11.25

From Theorem 11.5, you can see that two nonzero vectors meet at a right angle if and only if their dot product is zero. Two such vectors are said to be **orthogonal.**

Definition of Orthogonal Vectors

The vectors $\mathbf{u}$ and $\mathbf{v}$ are orthogonal if $\mathbf{u} \cdot \mathbf{v} = 0$.

NOTE The terms "perpendicular," "orthogonal," and "normal" all mean essentially the same thing—meeting at right angles. However, it is common to say that two vectors are *orthogonal*, two lines or planes are *perpendicular*, and a vector is *normal* to a given line or plane.

From this definition, it follows that the zero vector is orthogonal to every vector $\mathbf{u}$, because $\mathbf{0} \cdot \mathbf{u} = 0$. Moreover, for $0 \le \theta \le \pi$, you know that $\cos \theta = 0$ if and only if $\theta = \pi/2$. So, you can use Theorem 11.5 to conclude that two *nonzero* vectors are orthogonal if and only if the angle between them is $\pi/2$.

EXAMPLE 2 Finding the Angle Between Two Vectors

For $\mathbf{u} = \langle 3, -1, 2\rangle$, $\mathbf{v} = \langle -4, 0, 2\rangle$, $\mathbf{w} = \langle 1, -1, -2\rangle$, and $\mathbf{z} = \langle 2, 0, -1\rangle$, find the angle between each pair of vectors.

a. $\mathbf{u}$ and $\mathbf{v}$ **b.** $\mathbf{u}$ and $\mathbf{w}$ **c.** $\mathbf{v}$ and $\mathbf{z}$

Solution

a. $\cos \theta = \dfrac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\|\,\|\mathbf{v}\|} = \dfrac{-12 + 0 + 4}{\sqrt{14}\sqrt{20}} = \dfrac{-8}{2\sqrt{14}\sqrt{5}} = \dfrac{-4}{\sqrt{70}}$

Because $\mathbf{u} \cdot \mathbf{v} < 0$, $\theta = \arccos \dfrac{-4}{\sqrt{70}} \approx 2.069$ radians.

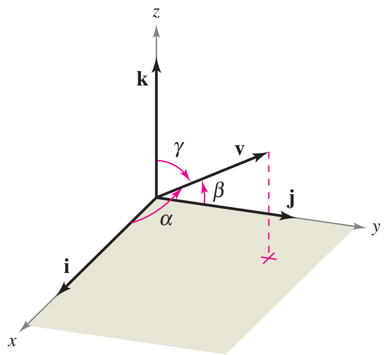
b. $\cos \theta = \dfrac{\mathbf{u} \cdot \mathbf{w}}{\|\mathbf{u}\|\,\|\mathbf{w}\|} = \dfrac{3 + 1 - 4}{\sqrt{14}\sqrt{6}} = \dfrac{0}{\sqrt{84}} = 0$

Because $\mathbf{u} \cdot \mathbf{w} = 0$, $\mathbf{u}$ and $\mathbf{w}$ are *orthogonal*. So, $\theta = \pi/2$.

c. $\cos \theta = \dfrac{\mathbf{v} \cdot \mathbf{z}}{\|\mathbf{v}\|\,\|\mathbf{z}\|} = \dfrac{-8 + 0 - 2}{\sqrt{20}\sqrt{5}} = \dfrac{-10}{\sqrt{100}} = -1$

Consequently, $\theta = \pi$. Note that $\mathbf{v}$ and $\mathbf{z}$ are parallel, with $\mathbf{v} = -2\mathbf{z}$.

Try It | Exploration A | Exploration B | Open Exploration



Direction angles
Figure 11.26

Direction Cosines

For a vector in the plane, you have seen that it is convenient to measure direction in terms of the angle, measured counterclockwise, *from* the positive x-axis *to* the vector. In space it is more convenient to measure direction in terms of the angles *between* the nonzero vector $\mathbf{v}$ and the three unit vectors $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$, as shown in Figure 11.26. The angles α, β, and γ are the **direction angles of v,** and $\cos \alpha$, $\cos \beta$, and $\cos \gamma$ are the **direction cosines of v.** Because

$$\mathbf{v} \cdot \mathbf{i} = \|\mathbf{v}\| \|\mathbf{i}\| \cos \alpha = \|\mathbf{v}\| \cos \alpha$$

and

$$\mathbf{v} \cdot \mathbf{i} = \langle v_1, v_2, v_3 \rangle \cdot \langle 1, 0, 0 \rangle = v_1$$

it follows that $\cos \alpha = v_1/\|\mathbf{v}\|$. By similar reasoning with the unit vectors $\mathbf{j}$ and $\mathbf{k}$, you have

$$\cos \alpha = \frac{v_1}{\|\mathbf{v}\|} \qquad \alpha \text{ is the angle between } \mathbf{v} \text{ and } \mathbf{i}.$$

$$\cos \beta = \frac{v_2}{\|\mathbf{v}\|} \qquad \beta \text{ is the angle between } \mathbf{v} \text{ and } \mathbf{j}.$$

$$\cos \gamma = \frac{v_3}{\|\mathbf{v}\|}. \qquad \gamma \text{ is the angle between } \mathbf{v} \text{ and } \mathbf{k}.$$

Consequently, any nonzero vector $\mathbf{v}$ in space has the normalized form

$$\frac{\mathbf{v}}{\|\mathbf{v}\|} = \frac{v_1}{\|\mathbf{v}\|}\mathbf{i} + \frac{v_2}{\|\mathbf{v}\|}\mathbf{j} + \frac{v_3}{\|\mathbf{v}\|}\mathbf{k} = \cos \alpha \mathbf{i} + \cos \beta \mathbf{j} + \cos \gamma \mathbf{k}$$

and because $\mathbf{v}/\|\mathbf{v}\|$ is a unit vector, it follows that

$$\cos^2 \alpha + \cos^2 \beta + \cos^2 \gamma = 1.$$

EXAMPLE 3 Finding Direction Angles

Find the direction cosines and angles for the vector $\mathbf{v} = 2\mathbf{i} + 3\mathbf{j} + 4\mathbf{k}$, and show that $\cos^2 \alpha + \cos^2 \beta + \cos^2 \gamma = 1$.

Solution Because $\|\mathbf{v}\| = \sqrt{2^2 + 3^2 + 4^2} = \sqrt{29}$, you can write the following.

$$\cos \alpha = \frac{v_1}{\|\mathbf{v}\|} = \frac{2}{\sqrt{29}} \Rightarrow \alpha \approx 68.2^\circ \qquad \text{Angle between } \mathbf{v} \text{ and } \mathbf{i}$$

$$\cos \beta = \frac{v_2}{\|\mathbf{v}\|} = \frac{3}{\sqrt{29}} \Rightarrow \beta \approx 56.1^\circ \qquad \text{Angle between } \mathbf{v} \text{ and } \mathbf{j}$$

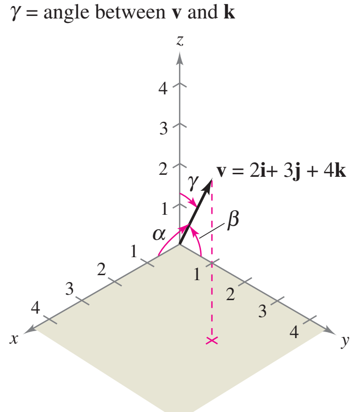
$$\cos \gamma = \frac{v_3}{\|\mathbf{v}\|} = \frac{4}{\sqrt{29}} \Rightarrow \gamma \approx 42.0^\circ \qquad \text{Angle between } \mathbf{v} \text{ and } \mathbf{k}$$

Furthermore, the sum of the squares of the direction cosines is

$$\begin{aligned} \cos^2 \alpha + \cos^2 \beta + \cos^2 \gamma &= \frac{4}{29} + \frac{9}{29} + \frac{16}{29} \\ &= \frac{29}{29} \\ &= 1. \end{aligned}$$

See Figure 11.27.

α = angle between $\mathbf{v}$ and $\mathbf{i}$
β = angle between $\mathbf{v}$ and $\mathbf{j}$
γ = angle between $\mathbf{v}$ and $\mathbf{k}$



The direction angles of $\mathbf{v}$
Figure 11.27

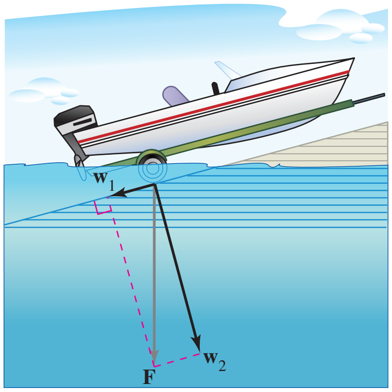
Try It | Exploration A | Exploration B | Exploration C

Projections and Vector Components

You have already seen applications in which two vectors are added to produce a resultant vector. Many applications in physics and engineering pose the reverse problem—decomposing a given vector into the sum of two **vector components.** The following physical example enables you to see the usefulness of this procedure.

Consider a boat on an inclined ramp, as shown in Figure 11.28. The force $\mathbf{F}$ due to gravity pulls the boat *down* the ramp and *against* the ramp. These two forces, $\mathbf{w}_1$ and $\mathbf{w}_2$, are orthogonal—they are called the vector components of $\mathbf{F}$.

$$\mathbf{F} = \mathbf{w}_1 + \mathbf{w}_2 \qquad \text{Vector components of } \mathbf{F}$$



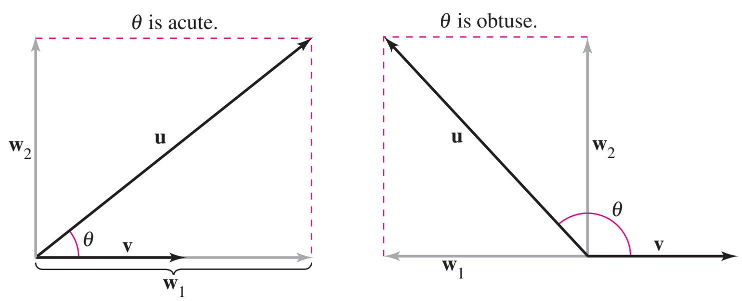
The force due to gravity pulls the boat against the ramp and down the ramp.
Figure 11.28

The forces $\mathbf{w}_1$ and $\mathbf{w}_2$ help you analyze the effect of gravity on the boat. For example, $\mathbf{w}_1$ indicates the force necessary to keep the boat from rolling down the ramp, whereas $\mathbf{w}_2$ indicates the force that the tires must withstand.

Definition of Projection and Vector Components

Let $\mathbf{u}$ and $\mathbf{v}$ be nonzero vectors. Moreover, let $\mathbf{u} = \mathbf{w}_1 + \mathbf{w}_2$, where $\mathbf{w}_1$ is parallel to $\mathbf{v}$ and $\mathbf{w}_2$ is orthogonal to $\mathbf{v}$, as shown in Figure 11.29.

1. $\mathbf{w}_1$ is called the **projection of u onto v** or the **vector component of u along v,** and is denoted by $\mathbf{w}_1 = \text{proj}_{\mathbf{v}}\mathbf{u}$.
2. $\mathbf{w}_2 = \mathbf{u} - \mathbf{w}_1$ is called the **vector component of u orthogonal to v.**



$\mathbf{w}_1 = \text{proj}_{\mathbf{v}}\mathbf{u} =$ projection of $\mathbf{u}$ onto $\mathbf{v}$ = vector component of $\mathbf{u}$ along $\mathbf{v}$
$\mathbf{w}_2 =$ vector component of $\mathbf{u}$ orthogonal to $\mathbf{v}$
Figure 11.29

EXAMPLE 4 Finding a Vector Component of u Orthogonal to v

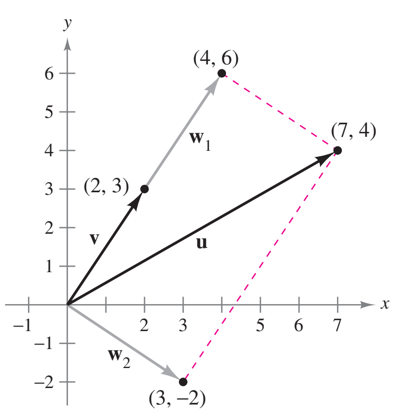
Find the vector component of $\mathbf{u} = \langle 7, 4\rangle$ that is orthogonal to $\mathbf{v} = \langle 2, 3\rangle$, given that $\mathbf{w}_1 = \text{proj}_{\mathbf{v}}\mathbf{u} = \langle 4, 6\rangle$ and

$$\mathbf{u} = \langle 7, 4\rangle = \mathbf{w}_1 + \mathbf{w}_2.$$

Solution Because $\mathbf{u} = \mathbf{w}_1 + \mathbf{w}_2$, where $\mathbf{w}_1$ is parallel to $\mathbf{v}$, it follows that $\mathbf{w}_2$ is the vector component of $\mathbf{u}$ orthogonal to $\mathbf{v}$. So, you have

$$\begin{aligned}\mathbf{w}_2 &= \mathbf{u} - \mathbf{w}_1\\ &= \langle 7, 4\rangle - \langle 4, 6\rangle\\ &= \langle 3, -2\rangle.\end{aligned}$$

Check to see that $\mathbf{w}_2$ is orthogonal to $\mathbf{v}$, as shown in Figure 11.30.



$\mathbf{u} = \mathbf{w}_1 + \mathbf{w}_2$
Figure 11.30

Try It

Exploration A

From Example 4, you can see that it is easy to find the vector component $\mathbf{w}_2$ once you have found the projection, $\mathbf{w}_1$, of $\mathbf{u}$ onto $\mathbf{v}$. To find this projection, use the dot product given in the theorem below, which you will prove in Exercise 90.

NOTE Note the distinction between the terms "component" and "vector component." For example, using the standard unit vectors with $\mathbf{u} = u_1\mathbf{i} + u_2\mathbf{j}$, u_1 is the *component* of $\mathbf{u}$ in the direction of $\mathbf{i}$ and $u_1\mathbf{i}$ is the *vector component* in the direction of $\mathbf{i}$.

THEOREM 11.6 Projection Using the Dot Product

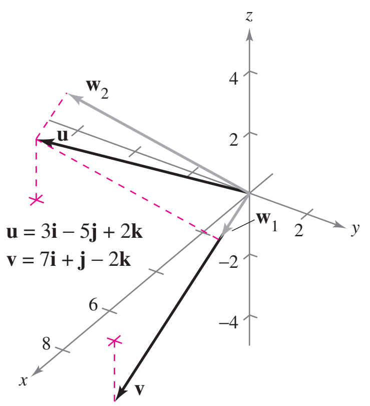
If $\mathbf{u}$ and $\mathbf{v}$ are nonzero vectors, then the projection of $\mathbf{u}$ onto $\mathbf{v}$ is given by

$$\text{proj}_{\mathbf{v}}\mathbf{u} = \left(\frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{v}\|^2}\right)\mathbf{v}.$$

The projection of $\mathbf{u}$ onto $\mathbf{v}$ can be written as a scalar multiple of a unit vector in the direction of $\mathbf{v}$. That is,

$$\left(\frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{v}\|^2}\right)\mathbf{v} = \left(\frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{v}\|}\right)\frac{\mathbf{v}}{\|\mathbf{v}\|} = (k)\frac{\mathbf{v}}{\|\mathbf{v}\|} \quad \Rightarrow \quad k = \frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{v}\|} = \|\mathbf{u}\|\cos\theta.$$

The scalar k is called the **component of u in the direction of v.**



$\mathbf{u} = \mathbf{w}_1 + \mathbf{w}_2$
Figure 11.31

EXAMPLE 5 Decomposing a Vector into Vector Components

Find the projection of $\mathbf{u}$ onto $\mathbf{v}$ and the vector component of $\mathbf{u}$ orthogonal to $\mathbf{v}$ for the vectors $\mathbf{u} = 3\mathbf{i} - 5\mathbf{j} + 2\mathbf{k}$ and $\mathbf{v} = 7\mathbf{i} + \mathbf{j} - 2\mathbf{k}$ shown in Figure 11.31.

Solution The projection of $\mathbf{u}$ onto $\mathbf{v}$ is

$$\mathbf{w}_1 = \left(\frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{v}\|^2}\right)\mathbf{v} = \left(\frac{12}{54}\right)(7\mathbf{i} + \mathbf{j} - 2\mathbf{k}) = \frac{14}{9}\mathbf{i} + \frac{2}{9}\mathbf{j} - \frac{4}{9}\mathbf{k}.$$

The vector component of $\mathbf{u}$ orthogonal to $\mathbf{v}$ is the vector

$$\mathbf{w}_2 = \mathbf{u} - \mathbf{w}_1 = (3\mathbf{i} - 5\mathbf{j} + 2\mathbf{k}) - \left(\frac{14}{9}\mathbf{i} + \frac{2}{9}\mathbf{j} - \frac{4}{9}\mathbf{k}\right) = \frac{13}{9}\mathbf{i} - \frac{47}{9}\mathbf{j} + \frac{22}{9}\mathbf{k}.$$

Try It Exploration A

EXAMPLE 6 Finding a Force

A 600-pound boat sits on a ramp inclined at 30°, as shown in Figure 11.32. What force is required to keep the boat from rolling down the ramp?

Solution Because the force due to gravity is vertical and downward, you can represent the gravitational force by the vector $\mathbf{F} = -600\mathbf{j}$. To find the force required to keep the boat from rolling down the ramp, project $\mathbf{F}$ onto a unit vector $\mathbf{v}$ in the direction of the ramp, as follows.

$$\mathbf{v} = \cos 30°\mathbf{i} + \sin 30°\mathbf{j} = \frac{\sqrt{3}}{2}\mathbf{i} + \frac{1}{2}\mathbf{j} \qquad \text{Unit vector along ramp}$$

Therefore, the projection of $\mathbf{F}$ onto $\mathbf{v}$ is given by

$$\mathbf{w}_1 = \text{proj}_{\mathbf{v}}\mathbf{F} = \left(\frac{\mathbf{F}\cdot\mathbf{v}}{\|\mathbf{v}\|^2}\right)\mathbf{v} = (\mathbf{F}\cdot\mathbf{v})\mathbf{v} = (-600)\left(\frac{1}{2}\right)\mathbf{v} = -300\left(\frac{\sqrt{3}}{2}\mathbf{i} + \frac{1}{2}\mathbf{j}\right).$$

The magnitude of this force is 300, and therefore a force of 300 pounds is required to keep the boat from rolling down the ramp.

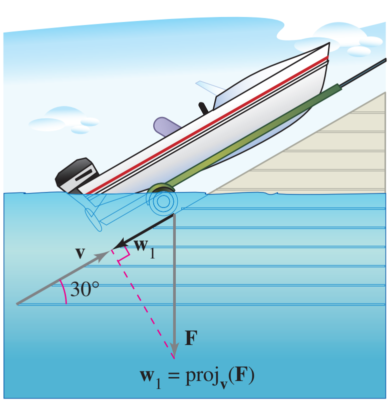
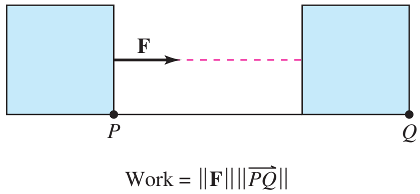


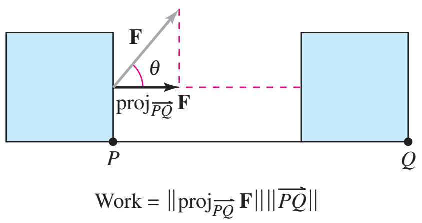
Figure 11.32

Simulation

Try It Exploration A



(a) Force acts along the line of motion.



(b) Force acts at angle θ with the line of motion.

Figure 11.33

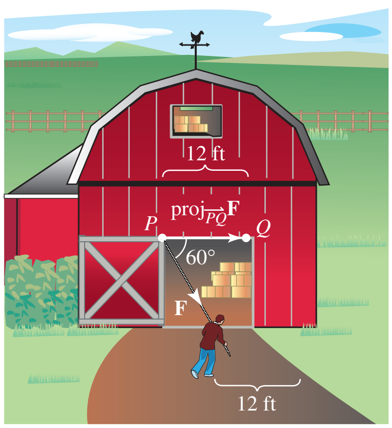


Figure 11.34

Work

The work W done by the constant force **F** acting along the line of motion of an object is given by

$$W = (\text{magnitude of force})(\text{distance}) = \|\mathbf{F}\|\,\|\overrightarrow{PQ}\|$$

as shown in Figure 11.33(a). If the constant force **F** is not directed along the line of motion, you can see from Figure 11.33(b) that the work W done by the force is

$$W = \|\text{proj}_{\overrightarrow{PQ}}\mathbf{F}\|\,\|\overrightarrow{PQ}\| = (\cos\theta)\|\mathbf{F}\|\,\|\overrightarrow{PQ}\| = \mathbf{F}\cdot\overrightarrow{PQ}.$$

This notion of work is summarized in the following definition.

Definition of Work

The work W done by a constant force **F** as its point of application moves along the vector $\overrightarrow{PQ}$ is given by either of the following.

1. $W = \|\text{proj}_{\overrightarrow{PQ}}\mathbf{F}\|\,\|\overrightarrow{PQ}\|$ Projection form
2. $W = \mathbf{F}\cdot\overrightarrow{PQ}$ Dot product form

EXAMPLE 7 Finding Work

To close a sliding door, a person pulls on a rope with a constant force of 50 pounds at a constant angle of 60°, as shown in Figure 11.34. Find the work done in moving the door 12 feet to its closed position.

Solution Using a projection, you can calculate the work as follows.

$$\begin{aligned} W &= \|\text{proj}_{\overrightarrow{PQ}}\mathbf{F}\|\,\|\overrightarrow{PQ}\| && \text{Projection form for work} \\ &= \cos(60^\circ)\,\|\mathbf{F}\|\,\|\overrightarrow{PQ}\| \\ &= \frac{1}{2}(50)(12) \\ &= 300 \text{ foot-pounds} \end{aligned}$$

Try It   Exploration A

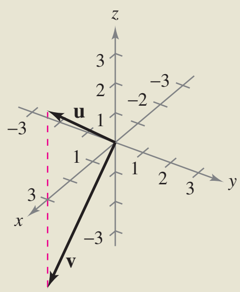
Section 11.4 The Cross Product of Two Vectors in Space

- **Find the cross product of two vectors in space.**
- **Use the triple scalar product of three vectors in space.**

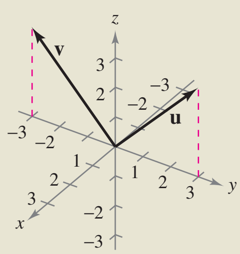
EXPLORATION

Geometric Property of the Cross Product Three pairs of vectors are shown below. Use the definition to find the cross product of each pair. Sketch all three vectors in a three-dimensional system. Describe any relationships among the three vectors. Use your description to write a conjecture about $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{u} \times \mathbf{v}$.

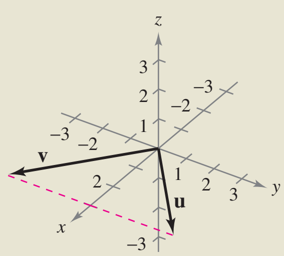
a. $\mathbf{u} = \langle 3, 0, 3\rangle$, $\mathbf{v} = \langle 3, 0, -3\rangle$



b. $\mathbf{u} = \langle 0, 3, 3\rangle$, $\mathbf{v} = \langle 0, -3, 3\rangle$



c. $\mathbf{u} = \langle 3, 3, 0\rangle$, $\mathbf{v} = \langle 3, -3, 0\rangle$



The Cross Product

Many applications in physics, engineering, and geometry involve finding a vector in space that is orthogonal to two given vectors. In this section you will study a product that will yield such a vector. It is called the **cross product,** and it is most conveniently defined and calculated using the standard unit vector form. Because the cross product yields a vector, it is also called the **vector product.**

Definition of Cross Product of Two Vectors in Space

Let $\mathbf{u} = u_1\mathbf{i} + u_2\mathbf{j} + u_3\mathbf{k}$ and $\mathbf{v} = v_1\mathbf{i} + v_2\mathbf{j} + v_3\mathbf{k}$ be vectors in space. The **cross product** of $\mathbf{u}$ and $\mathbf{v}$ is the vector

$$\mathbf{u} \times \mathbf{v} = (u_2v_3 - u_3v_2)\mathbf{i} - (u_1v_3 - u_3v_1)\mathbf{j} + (u_1v_2 - u_2v_1)\mathbf{k}.$$

NOTE Be sure you see that this definition applies only to three-dimensional vectors. The cross product is not defined for two-dimensional vectors.

A convenient way to calculate $\mathbf{u} \times \mathbf{v}$ is to use the following *determinant form* with cofactor expansion. (This 3×3 determinant form is used simply to help remember the formula for the cross product—it is technically not a determinant because the entries of the corresponding matrix are not all real numbers.)

$$\mathbf{u} \times \mathbf{v} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix} \begin{array}{l} \\ \longleftarrow \text{Put “}\mathbf{u}\text{” in Row 2.} \\ \longleftarrow \text{Put “}\mathbf{v}\text{” in Row 3.} \end{array}$$

$$= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix}\mathbf{i} - \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix}\mathbf{j} + \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix}\mathbf{k}$$

$$= \begin{vmatrix} u_2 & u_3 \\ v_2 & v_3 \end{vmatrix}\mathbf{i} - \begin{vmatrix} u_1 & u_3 \\ v_1 & v_3 \end{vmatrix}\mathbf{j} + \begin{vmatrix} u_1 & u_2 \\ v_1 & v_2 \end{vmatrix}\mathbf{k}$$

$$= (u_2v_3 - u_3v_2)\mathbf{i} - (u_1v_3 - u_3v_1)\mathbf{j} + (u_1v_2 - u_2v_1)\mathbf{k}$$

Note the minus sign in front of the $\mathbf{j}$-component. Each of the three 2×2 determinants can be evaluated by using the following diagonal pattern.

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

Here are a couple of examples.

$$\begin{vmatrix} 2 & 4 \\ 3 & -1 \end{vmatrix} = (2)(-1) - (4)(3) = -2 - 12 = -14$$

$$\begin{vmatrix} 4 & 0 \\ -6 & 3 \end{vmatrix} = (4)(3) - (0)(-6) = 12$$

Notation for Dot and Cross Products

The notation for the dot product and cross product of vectors was first introduced by the American physicist Josiah Willard Gibbs (1839–1903). In the early 1880s, Gibbs built a system to represent physical quantities called "vector analysis." The system was a departure from Hamilton's theory of quaternions.

MathBio

EXAMPLE 1 Finding the Cross Product

Given $\mathbf{u} = \mathbf{i} - 2\mathbf{j} + \mathbf{k}$ and $\mathbf{v} = 3\mathbf{i} + \mathbf{j} - 2\mathbf{k}$, find each of the following.

a. $\mathbf{u} \times \mathbf{v}$ **b.** $\mathbf{v} \times \mathbf{u}$ **c.** $\mathbf{v} \times \mathbf{v}$

Solution

$$\textbf{a. } \mathbf{u} \times \mathbf{v} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & -2 & 1 \\ 3 & 1 & -2 \end{vmatrix} = \begin{vmatrix} -2 & 1 \\ 1 & -2 \end{vmatrix}\mathbf{i} - \begin{vmatrix} 1 & 1 \\ 3 & -2 \end{vmatrix}\mathbf{j} + \begin{vmatrix} 1 & -2 \\ 3 & 1 \end{vmatrix}\mathbf{k}$$

$$= (4 - 1)\mathbf{i} - (-2 - 3)\mathbf{j} + (1 + 6)\mathbf{k}$$

$$= 3\mathbf{i} + 5\mathbf{j} + 7\mathbf{k}$$

$$\textbf{b. } \mathbf{v} \times \mathbf{u} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 3 & 1 & -2 \\ 1 & -2 & 1 \end{vmatrix} = \begin{vmatrix} 1 & -2 \\ -2 & 1 \end{vmatrix}\mathbf{i} - \begin{vmatrix} 3 & -2 \\ 1 & 1 \end{vmatrix}\mathbf{j} + \begin{vmatrix} 3 & 1 \\ 1 & -2 \end{vmatrix}\mathbf{k}$$

$$= (1 - 4)\mathbf{i} - (3 + 2)\mathbf{j} + (-6 - 1)\mathbf{k}$$

$$= -3\mathbf{i} - 5\mathbf{j} - 7\mathbf{k}$$

Note that this result is the negative of that in part (a).

$$\textbf{c. } \mathbf{v} \times \mathbf{v} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 3 & 1 & -2 \\ 3 & 1 & -2 \end{vmatrix} = \mathbf{0}$$

Try It Exploration A Exploration B Exploration C

The results obtained in Example 1 suggest some interesting *algebraic* properties of the cross product. For instance, $\mathbf{u} \times \mathbf{v} = -(\mathbf{v} \times \mathbf{u})$, and $\mathbf{v} \times \mathbf{v} = \mathbf{0}$. These properties, and several others, are summarized in the following theorem.

THEOREM 11.7 Algebraic Properties of the Cross Product

Let $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ be vectors in space, and let c be a scalar.

1. $\mathbf{u} \times \mathbf{v} = -(\mathbf{v} \times \mathbf{u})$
2. $\mathbf{u} \times (\mathbf{v} + \mathbf{w}) = (\mathbf{u} \times \mathbf{v}) + (\mathbf{u} \times \mathbf{w})$
3. $c(\mathbf{u} \times \mathbf{v}) = (c\mathbf{u}) \times \mathbf{v} = \mathbf{u} \times (c\mathbf{v})$
4. $\mathbf{u} \times \mathbf{0} = \mathbf{0} \times \mathbf{u} = \mathbf{0}$
5. $\mathbf{u} \times \mathbf{u} = \mathbf{0}$
6. $\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) = (\mathbf{u} \times \mathbf{v}) \cdot \mathbf{w}$

Proof To prove Property 1, let $\mathbf{u} = u_1\mathbf{i} + u_2\mathbf{j} + u_3\mathbf{k}$ and $\mathbf{v} = v_1\mathbf{i} + v_2\mathbf{j} + v_3\mathbf{k}$. Then,

$$\mathbf{u} \times \mathbf{v} = (u_2v_3 - u_3v_2)\mathbf{i} - (u_1v_3 - u_3v_1)\mathbf{j} + (u_1v_2 - u_2v_1)\mathbf{k}$$

and

$$\mathbf{v} \times \mathbf{u} = (v_2u_3 - v_3u_2)\mathbf{i} - (v_1u_3 - v_3u_1)\mathbf{j} + (v_1u_2 - v_2u_1)\mathbf{k}$$

which implies that $\mathbf{u} \times \mathbf{v} = -(\mathbf{v} \times \mathbf{u})$. Proofs of Properties 2, 3, 5, and 6 are left as exercises (see Exercises 57–60).

Note that Property 1 of Theorem 11.7 indicates that the cross product is *not commutative*. In particular, this property indicates that the vectors $\mathbf{u} \times \mathbf{v}$ and $\mathbf{v} \times \mathbf{u}$ have equal lengths but opposite directions. The following theorem lists some other *geometric* properties of the cross product of two vectors.

NOTE It follows from Properties 1 and 2 in Theorem 11.8 that if $\mathbf{n}$ is a unit vector orthogonal to both $\mathbf{u}$ and $\mathbf{v}$, then

$$\mathbf{u} \times \mathbf{v} = \pm(\|\mathbf{u}\|\ \|\mathbf{v}\| \sin\theta)\mathbf{n}.$$

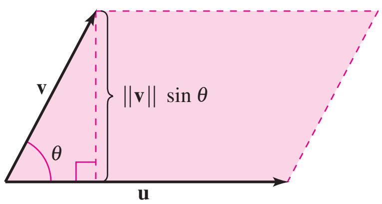
THEOREM 11.8 Geometric Properties of the Cross Product

Let $\mathbf{u}$ and $\mathbf{v}$ be nonzero vectors in space, and let θ be the angle between $\mathbf{u}$ and $\mathbf{v}$.

1. $\mathbf{u} \times \mathbf{v}$ is orthogonal to both $\mathbf{u}$ and $\mathbf{v}$.
2. $\|\mathbf{u} \times \mathbf{v}\| = \|\mathbf{u}\|\ \|\mathbf{v}\| \sin\theta$
3. $\mathbf{u} \times \mathbf{v} = \mathbf{0}$ if and only if $\mathbf{u}$ and $\mathbf{v}$ are scalar multiples of each other.
4. $\|\mathbf{u} \times \mathbf{v}\|$ = area of parallelogram having $\mathbf{u}$ and $\mathbf{v}$ as adjacent sides.

Proof To prove Property 2, note because $\cos\theta = (\mathbf{u} \cdot \mathbf{v})/(\|\mathbf{u}\|\,\|\mathbf{v}\|)$, it follows that

$$\begin{aligned}
\|\mathbf{u}\|\ \|\mathbf{v}\| \sin\theta &= \|\mathbf{u}\|\ \|\mathbf{v}\|\sqrt{1-\cos^2\theta} \\
&= \|\mathbf{u}\|\ \|\mathbf{v}\|\sqrt{1-\frac{(\mathbf{u}\cdot\mathbf{v})^2}{\|\mathbf{u}\|^2\|\mathbf{v}\|^2}} \\
&= \sqrt{\|\mathbf{u}\|^2\|\mathbf{v}\|^2-(\mathbf{u}\cdot\mathbf{v})^2} \\
&= \sqrt{(u_1^2+u_2^2+u_3^2)(v_1^2+v_2^2+v_3^2)-(u_1v_1+u_2v_2+u_3v_3)^2} \\
&= \sqrt{(u_2v_3-u_3v_2)^2+(u_1v_3-u_3v_1)^2+(u_1v_2-u_2v_1)^2} \\
&= \|\mathbf{u}\times\mathbf{v}\|.
\end{aligned}$$



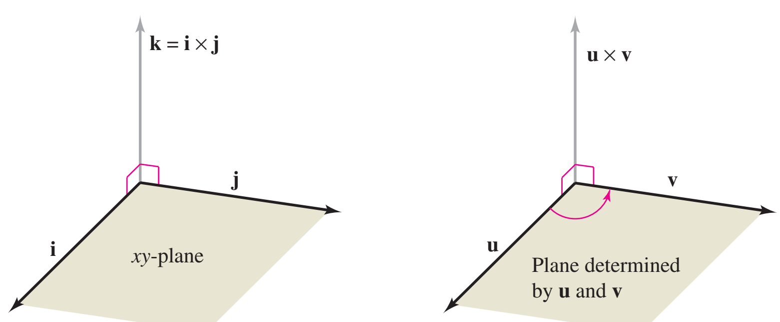
The vectors $\mathbf{u}$ and $\mathbf{v}$ form adjacent sides of a parallelogram.
Figure 11.35

To prove Property 4, refer to Figure 11.35, which is a parallelogram having $\mathbf{v}$ and $\mathbf{u}$ as adjacent sides. Because the height of the parallelogram is $\|\mathbf{v}\| \sin\theta$, the area is

$$\begin{aligned}
\text{Area} &= (\text{base})(\text{height}) \\
&= \|\mathbf{u}\|\ \|\mathbf{v}\| \sin\theta \\
&= \|\mathbf{u}\times\mathbf{v}\|.
\end{aligned}$$

Proofs of Properties 1 and 3 are left as exercises (see Exercises 61 and 62).

Both $\mathbf{u} \times \mathbf{v}$ and $\mathbf{v} \times \mathbf{u}$ are perpendicular to the plane determined by $\mathbf{u}$ and $\mathbf{v}$. One way to remember the orientations of the vectors $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{u} \times \mathbf{v}$ is to compare them with the unit vectors $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k} = \mathbf{i} \times \mathbf{j}$, as shown in Figure 11.36. The three vectors $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{u} \times \mathbf{v}$ form a *right-handed system*, whereas the three vectors $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{v} \times \mathbf{u}$ form a *left-handed system*.



Right-handed systems
Figure 11.36

EXAMPLE 2 Using the Cross Product

Find a unit vector that is orthogonal to both

$$\mathbf{u} = \mathbf{i} - 4\mathbf{j} + \mathbf{k} \quad \text{and} \quad \mathbf{v} = 2\mathbf{i} + 3\mathbf{j}.$$

Solution The cross product $\mathbf{u} \times \mathbf{v}$, as shown in Figure 11.37, is orthogonal to both $\mathbf{u}$ and $\mathbf{v}$.

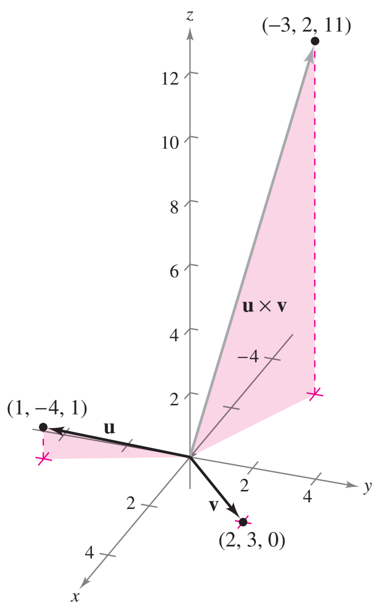
$$\begin{aligned}\mathbf{u} \times \mathbf{v} &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & -4 & 1 \\ 2 & 3 & 0 \end{vmatrix} \qquad \text{Cross product} \\ &= -3\mathbf{i} + 2\mathbf{j} + 11\mathbf{k}\end{aligned}$$

Because

$$\|\mathbf{u} \times \mathbf{v}\| = \sqrt{(-3)^2 + 2^2 + 11^2} = \sqrt{134}$$

a unit vector orthogonal to both $\mathbf{u}$ and $\mathbf{v}$ is

$$\frac{\mathbf{u} \times \mathbf{v}}{\|\mathbf{u} \times \mathbf{v}\|} = -\frac{3}{\sqrt{134}}\mathbf{i} + \frac{2}{\sqrt{134}}\mathbf{j} + \frac{11}{\sqrt{134}}\mathbf{k}.$$



The vector $\mathbf{u} \times \mathbf{v}$ is orthogonal to both $\mathbf{u}$ and $\mathbf{v}$.
Figure 11.37

Try It | Exploration A | Exploration B | Open Exploration

NOTE In Example 2, note that you could have used the cross product $\mathbf{v} \times \mathbf{u}$ to form a unit vector that is orthogonal to both $\mathbf{u}$ and $\mathbf{v}$. With that choice, you would have obtained the negative of the unit vector found in the example.

EXAMPLE 3 Geometric Application of the Cross Product

Show that the quadrilateral with vertices at the following points is a parallelogram, and find its area.

$$A = (5, 2, 0) \qquad B = (2, 6, 1)$$
$$C = (2, 4, 7) \qquad D = (5, 0, 6)$$

Solution From Figure 11.38 you can see that the sides of the quadrilateral correspond to the following four vectors.

$$\overrightarrow{AB} = -3\mathbf{i} + 4\mathbf{j} + \mathbf{k} \qquad \overrightarrow{CD} = 3\mathbf{i} - 4\mathbf{j} - \mathbf{k} = -\overrightarrow{AB}$$
$$\overrightarrow{AD} = 0\mathbf{i} - 2\mathbf{j} + 6\mathbf{k} \qquad \overrightarrow{CB} = 0\mathbf{i} + 2\mathbf{j} - 6\mathbf{k} = -\overrightarrow{AD}$$

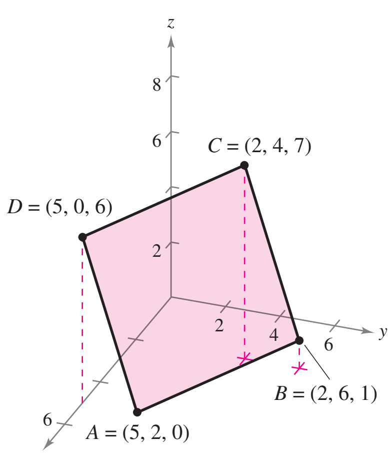
So, $\overrightarrow{AB}$ is parallel to $\overrightarrow{CD}$ and $\overrightarrow{AD}$ is parallel to $\overrightarrow{CB}$, and you can conclude that the quadrilateral is a parallelogram with $\overrightarrow{AB}$ and $\overrightarrow{AD}$ as adjacent sides. Moreover, because

$$\begin{aligned}\overrightarrow{AB} \times \overrightarrow{AD} &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ -3 & 4 & 1 \\ 0 & -2 & 6 \end{vmatrix} \qquad \text{Cross product} \\ &= 26\mathbf{i} + 18\mathbf{j} + 6\mathbf{k}\end{aligned}$$

the area of the parallelogram is

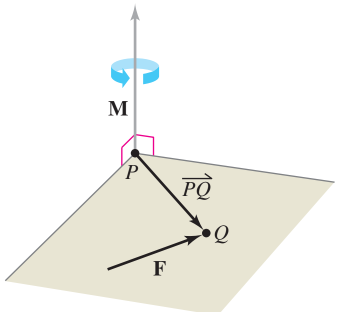
$$\|\overrightarrow{AB} \times \overrightarrow{AD}\| = \sqrt{1036} \approx 32.19.$$

Is the parallelogram a rectangle? You can determine whether it is by finding the angle between the vectors $\overrightarrow{AB}$ and $\overrightarrow{AD}$.



The area of the parallelogram is approximately 32.19.
Figure 11.38

Try It | Exploration A



The moment of **F** about P
Figure 11.39

In physics, the cross product can be used to measure **torque**—the **moment M of a force F about a point *P*,** as shown in Figure 11.39. If the point of application of the force is Q, the moment of **F** about P is given by

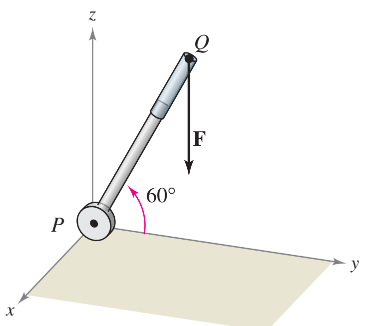
$$\mathbf{M} = \overrightarrow{PQ} \times \mathbf{F}.$$

Moment of **F** about P

The magnitude of the moment **M** measures the tendency of the vector $\overrightarrow{PQ}$ to rotate counterclockwise (using the right-hand rule) about an axis directed along the vector **M**.

EXAMPLE 4 An Application of the Cross Product

A vertical force of 50 pounds is applied to the end of a one-foot lever that is attached to an axle at point P, as shown in Figure 11.40. Find the moment of this force about the point P when $\theta = 60°$.



A vertical force of 50 pounds is applied at point Q.
Figure 11.40

Solution If you represent the 50-pound force as $\mathbf{F} = -50\mathbf{k}$ and the lever as

$$\overrightarrow{PQ} = \cos(60°)\mathbf{j} + \sin(60°)\mathbf{k} = \frac{1}{2}\mathbf{j} + \frac{\sqrt{3}}{2}\mathbf{k}$$

the moment of **F** about P is given by

$$\mathbf{M} = \overrightarrow{PQ} \times \mathbf{F} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 0 & \frac{1}{2} & \frac{\sqrt{3}}{2} \\ 0 & 0 & -50 \end{vmatrix} = -25\mathbf{i}.$$

Moment of **F** about P

The magnitude of this moment is 25 foot-pounds.

Try It Exploration A

NOTE In Example 4, note that the moment (the tendency of the lever to rotate about its axle) is dependent on the angle θ. When $\theta = \pi/2$, the moment is $\mathbf{0}$. The moment is greatest when $\theta = 0$.

The Triple Scalar Product

For vectors **u**, **v**, and **w** in space, the dot product of **u** and $\mathbf{v} \times \mathbf{w}$

$$\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w})$$

is called the **triple scalar product,** as defined in Theorem 11.9. The proof of this theorem is left as an exercise (see Exercise 56).

FOR FURTHER INFORMATION To see how the cross product is used to model the torque of the robot arm of a space shuttle, see the article "The Long Arm of Calculus" by Ethan Berkove and Rich Marchand in *The College Mathematics Journal*.

MathArticle

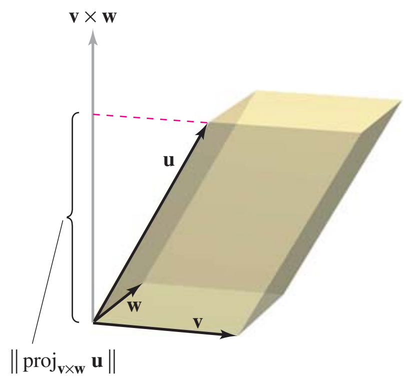
THEOREM 11.9 The Triple Scalar Product

For $\mathbf{u} = u_1\mathbf{i} + u_2\mathbf{j} + u_3\mathbf{k}$, $\mathbf{v} = v_1\mathbf{i} + v_2\mathbf{j} + v_3\mathbf{k}$, and $\mathbf{w} = w_1\mathbf{i} + w_2\mathbf{j} + w_3\mathbf{k}$, the triple scalar product is given by

$$\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) = \begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix}.$$

NOTE The value of a determinant is multiplied by -1 if two rows are interchanged. After two such interchanges, the value of the determinant will be unchanged. So, the following triple scalar products are equivalent.

$$\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) = \mathbf{v} \cdot (\mathbf{w} \times \mathbf{u}) = \mathbf{w} \cdot (\mathbf{u} \times \mathbf{v})$$



Area of base $= \|\mathbf{v} \times \mathbf{w}\|$
Volume of parallelepiped $= |\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w})|$
Figure 11.41

Rotatable Graph

If the vectors **u**, **v**, and **w** do not lie in the same plane, the triple scalar product $\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w})$ can be used to determine the volume of the parallelepiped (a polyhedron, all of whose faces are parallelograms) with **u**, **v**, and **w** as adjacent edges, as shown in Figure 11.41. This is established in the following theorem.

THEOREM 11.10 Geometric Property of Triple Scalar Product

The volume V of a parallelepiped with vectors **u**, **v**, and **w** as adjacent edges is given by

$$V = |\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w})|.$$

Proof In Figure 11.41, note that

$$\|\mathbf{v} \times \mathbf{w}\| = \text{area of base}$$

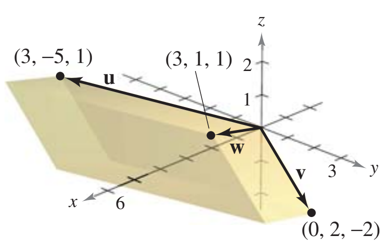
and

$$\|\text{proj}_{\mathbf{v}\times\mathbf{w}}\mathbf{u}\| = \text{height of parallelepiped.}$$

Therefore, the volume is

$$\begin{aligned} V = (\text{height})(\text{area of base}) &= \|\text{proj}_{\mathbf{v}\times\mathbf{w}}\mathbf{u}\|\,\|\mathbf{v} \times \mathbf{w}\| \\ &= \left|\frac{\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w})}{\|\mathbf{v} \times \mathbf{w}\|}\right| \|\mathbf{v} \times \mathbf{w}\| \\ &= |\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w})|. \end{aligned}$$

EXAMPLE 5 Volume by the Triple Scalar Product



The parallelepiped has a volume of 36.
Figure 11.42

Rotatable Graph

Find the volume of the parallelepiped shown in Figure 11.42 having $\mathbf{u} = 3\mathbf{i} - 5\mathbf{j} + \mathbf{k}$, $\mathbf{v} = 2\mathbf{j} - 2\mathbf{k}$, and $\mathbf{w} = 3\mathbf{i} + \mathbf{j} + \mathbf{k}$ as adjacent edges.

Solution By Theorem 11.10, you have

$$\begin{aligned} V &= |\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w})| \qquad \text{Triple scalar product} \\ &= \begin{vmatrix} 3 & -5 & 1 \\ 0 & 2 & -2 \\ 3 & 1 & 1 \end{vmatrix} \\ &= 3\begin{vmatrix} 2 & -2 \\ 1 & 1 \end{vmatrix} - (-5)\begin{vmatrix} 0 & -2 \\ 3 & 1 \end{vmatrix} + (1)\begin{vmatrix} 0 & 2 \\ 3 & 1 \end{vmatrix} \\ &= 3(4) + 5(6) + 1(-6) \\ &= 36. \end{aligned}$$

Try It | Exploration A | Exploration B

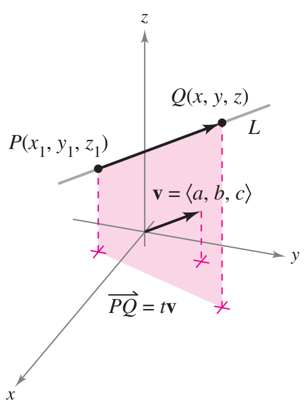
A natural consequence of Theorem 11.10 is that the volume of the parallelepiped is 0 if and only if the three vectors are coplanar. That is, if the vectors $\mathbf{u} = \langle u_1, u_2, u_3 \rangle$, $\mathbf{v} = \langle v_1, v_2, v_3 \rangle$, and $\mathbf{w} = \langle w_1, w_2, w_3 \rangle$ have the same initial point, they lie in the same plane if and only if

$$\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) = \begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix} = 0.$$

Section 11.5 Lines and Planes in Space

- **Write a set of parametric equations for a line in space.**
- **Write a linear equation to represent a plane in space.**
- **Sketch the plane given by a linear equation.**
- **Find the distances between points, planes, and lines in space.**

Lines in Space



Line L and its direction vector $\mathbf{v}$
Figure 11.43

In the plane, *slope* is used to determine an equation of a line. In space, it is more convenient to use *vectors* to determine the equation of a line.

In Figure 11.43, consider the line L through the point $P(x_1, y_1, z_1)$ and parallel to the vector $\mathbf{v} = \langle a, b, c \rangle$. The vector $\mathbf{v}$ is a **direction vector** for the line L, and a, b, and c are **direction numbers.** One way of describing the line L is to say that it consists of all points $Q(x, y, z)$ for which the vector $\overrightarrow{PQ}$ is parallel to $\mathbf{v}$. This means that $\overrightarrow{PQ}$ is a scalar multiple of $\mathbf{v}$, and you can write $\overrightarrow{PQ} = t\mathbf{v}$, where t is a scalar (a real number).

$$\overrightarrow{PQ} = \langle x - x_1, y - y_1, z - z_1 \rangle = \langle at, bt, ct \rangle = t\mathbf{v}$$

By equating corresponding components, you can obtain **parametric equations** of a line in space.

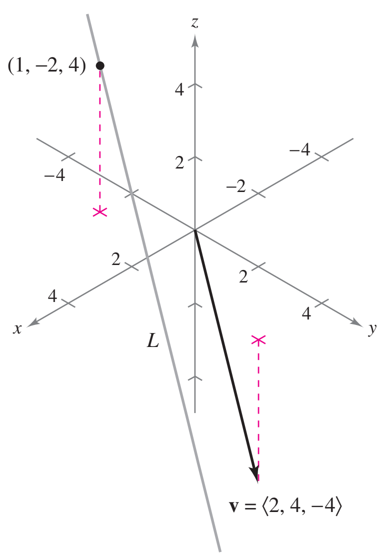
THEOREM 11.11 Parametric Equations of a Line in Space

A line L parallel to the vector $\mathbf{v} = \langle a, b, c \rangle$ and passing through the point $P(x_1, y_1, z_1)$ is represented by the **parametric equations**

$$x = x_1 + at, \quad y = y_1 + bt, \quad \text{and} \quad z = z_1 + ct.$$

If the direction numbers a, b, and c are all nonzero, you can eliminate the parameter t to obtain **symmetric equations** of the line.

$$\frac{x - x_1}{a} = \frac{y - y_1}{b} = \frac{z - z_1}{c} \qquad \text{Symmetric equations}$$



The vector $\mathbf{v}$ is parallel to the line L.
Figure 11.44

EXAMPLE 1 Finding Parametric and Symmetric Equations

Find parametric and symmetric equations of the line L that passes through the point $(1, -2, 4)$ and is parallel to $\mathbf{v} = \langle 2, 4, -4 \rangle$.

Solution To find a set of parametric equations of the line, use the coordinates $x_1 = 1$, $y_1 = -2$, and $z_1 = 4$ and direction numbers $a = 2$, $b = 4$, and $c = -4$ (see Figure 11.44).

$$x = 1 + 2t, \quad y = -2 + 4t, \quad z = 4 - 4t \qquad \text{Parametric equations}$$

Because a, b, and c are all nonzero, a set of symmetric equations is

$$\frac{x - 1}{2} = \frac{y + 2}{4} = \frac{z - 4}{-4}. \qquad \text{Symmetric equations}$$

Try It   Exploration A

Neither parametric equations nor symmetric equations of a given line are unique. For instance, in Example 1, by letting $t = 1$ in the parametric equations you would obtain the point $(3, 2, 0)$. Using this point with the direction numbers $a = 2$, $b = 4$, and $c = -4$ would produce a different set of parametric equations

$$x = 3 + 2t, \quad y = 2 + 4t, \quad \text{and} \quad z = -4t.$$

EXAMPLE 2 Parametric Equations of a Line Through Two Points

Find a set of parametric equations of the line that passes through the points $(-2, 1, 0)$ and $(1, 3, 5)$.

Solution Begin by using the points $P(-2, 1, 0)$ and $Q(1, 3, 5)$ to find a direction vector for the line passing through P and Q, given by

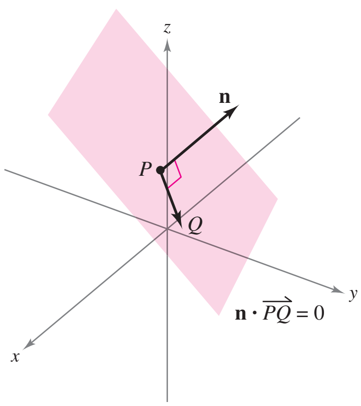
$$\mathbf{v} = \overrightarrow{PQ} = \langle 1 - (-2), 3 - 1, 5 - 0 \rangle = \langle 3, 2, 5 \rangle = \langle a, b, c \rangle.$$

Using the direction numbers $a = 3$, $b = 2$, and $c = 5$ with the point $P(-2, 1, 0)$, you can obtain the parametric equations

$$x = -2 + 3t, \quad y = 1 + 2t, \quad \text{and} \quad z = 5t.$$

Try It Exploration A Open Exploration

NOTE As t varies over all real numbers, the parametric equations in Example 2 determine the points (x, y, z) on the line. In particular, note that $t = 0$ and $t = 1$ give the original points $(-2, 1, 0)$ and $(1, 3, 5)$.



The normal vector $\mathbf{n}$ is orthogonal to each vector $\overrightarrow{PQ}$ in the plane.
Figure 11.45

Planes in Space

You have seen how an equation of a line in space can be obtained from a point on the line and a vector *parallel* to it. You will now see that an equation of a plane in space can be obtained from a point in the plane and a vector *normal* (perpendicular) to the plane.

Consider the plane containing the point $P(x_1, y_1, z_1)$ having a nonzero normal vector $\mathbf{n} = \langle a, b, c \rangle$, as shown in Figure 11.45. This plane consists of all points $Q(x, y, z)$ for which vector $\overrightarrow{PQ}$ is orthogonal to $\mathbf{n}$. Using the dot product, you can write the following.

$$\begin{aligned} \mathbf{n} \cdot \overrightarrow{PQ} &= 0 \\ \langle a, b, c \rangle \cdot \langle x - x_1, y - y_1, z - z_1 \rangle &= 0 \\ a(x - x_1) + b(y - y_1) + c(z - z_1) &= 0 \end{aligned}$$

The third equation of the plane is said to be in **standard form.**

THEOREM 11.12 Standard Equation of a Plane in Space

The plane containing the point (x_1, y_1, z_1) and having a normal vector $\mathbf{n} = \langle a, b, c \rangle$ can be represented, in **standard form,** by the equation

$$a(x - x_1) + b(y - y_1) + c(z - z_1) = 0.$$

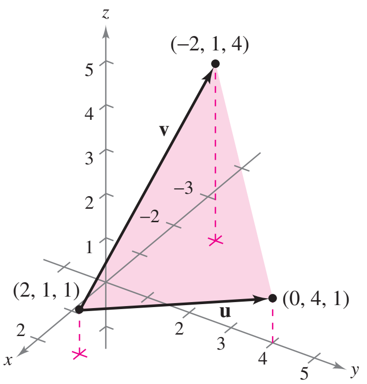
By regrouping terms, you obtain the **general form** of the equation of a plane in space.

$$ax + by + cz + d = 0$$ General form of equation of plane

Given the general form of the equation of a plane, it is easy to find a normal vector to the plane. Simply use the coefficients of x, y, and z and write $\mathbf{n} = \langle a, b, c\rangle$.

EXAMPLE 3 Finding an Equation of a Plane in Three-Space

Find the general equation of the plane containing the points $(2, 1, 1)$, $(0, 4, 1)$, and $(-2, 1, 4)$.



A plane determined by **u** and **v**
Figure 11.46

Solution To apply Theorem 11.12 you need a point in the plane and a vector that is normal to the plane. There are three choices for the point, but no normal vector is given. To obtain a normal vector, use the cross product of vectors **u** and **v** extending from the point $(2, 1, 1)$ to the points $(0, 4, 1)$ and $(-2, 1, 4)$, as shown in Figure 11.46. The component forms of **u** and **v** are

$$\mathbf{u} = \langle 0 - 2, 4 - 1, 1 - 1\rangle = \langle -2, 3, 0\rangle$$
$$\mathbf{v} = \langle -2 - 2, 1 - 1, 4 - 1\rangle = \langle -4, 0, 3\rangle$$

and it follows that

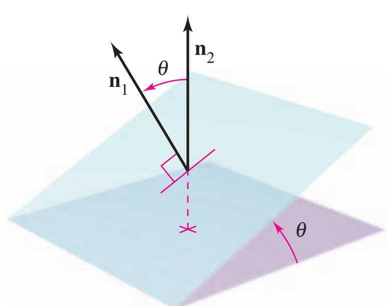
$$\begin{aligned}\mathbf{n} &= \mathbf{u} \times \mathbf{v}\\ &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ -2 & 3 & 0 \\ -4 & 0 & 3 \end{vmatrix}\\ &= 9\mathbf{i} + 6\mathbf{j} + 12\mathbf{k}\\ &= \langle a, b, c\rangle\end{aligned}$$

is normal to the given plane. Using the direction numbers for **n** and the point $(x_1, y_1, z_1) = (2, 1, 1)$, you can determine an equation of the plane to be

$$a(x - x_1) + b(y - y_1) + c(z - z_1) = 0$$
$$9(x - 2) + 6(y - 1) + 12(z - 1) = 0 \qquad \text{Standard form}$$
$$9x + 6y + 12z - 36 = 0 \qquad \text{General form}$$
$$3x + 2y + 4z - 12 = 0. \qquad \text{Simplified general form}$$

Try It | Exploration A | Exploration B

NOTE In Example 3, check to see that each of the three original points satisfies the equation $3x + 2y + 4z - 12 = 0$.



The angle θ between two planes
Figure 11.47

Rotatable Graph

Two distinct planes in three-space either are parallel or intersect in a line. If they intersect, you can determine the angle $(0 \le \theta \le \pi/2)$ between them from the angle between their normal vectors, as shown in Figure 11.47. Specifically, if vectors $\mathbf{n}_1$ and $\mathbf{n}_2$ are normal to two intersecting planes, the angle θ between the normal vectors is equal to the angle between the two planes and is given by

$$\cos\theta = \frac{|\mathbf{n}_1 \cdot \mathbf{n}_2|}{\|\mathbf{n}_1\|\,\|\mathbf{n}_2\|}. \qquad \text{Angle between two planes}$$

Consequently, two planes with normal vectors $\mathbf{n}_1$ and $\mathbf{n}_2$ are

1. *perpendicular* if $\mathbf{n}_1 \cdot \mathbf{n}_2 = 0$.
2. *parallel* if $\mathbf{n}_1$ is a scalar multiple of $\mathbf{n}_2$.

EXAMPLE 4 Finding the Line of Intersection of Two Planes

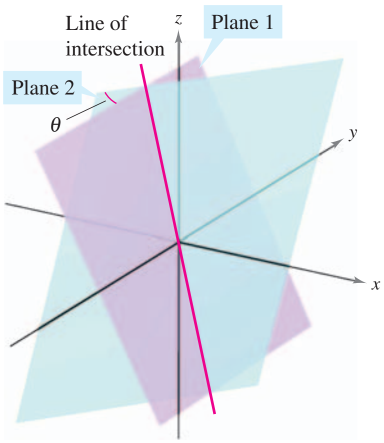


Figure 11.48

Rotatable Graph

Find the angle between the two planes given by

$$x - 2y + z = 0 \qquad \text{Equation of plane 1}$$
$$2x + 3y - 2z = 0 \qquad \text{Equation of plane 2}$$

and find parametric equations of their line of intersection (see Figure 11.48).

Solution Normal vectors for the planes are $\mathbf{n}_1 = \langle 1, -2, 1\rangle$ and $\mathbf{n}_2 = \langle 2, 3, -2\rangle$. Consequently, the angle between the two planes is determined as follows.

$$\begin{aligned}\cos\theta &= \frac{|\mathbf{n}_1 \cdot \mathbf{n}_2|}{\|\mathbf{n}_1\|\,\|\mathbf{n}_2\|} \qquad \text{Cosine of angle between } \mathbf{n}_1 \text{ and } \mathbf{n}_2\\ &= \frac{|-6|}{\sqrt{6}\sqrt{17}}\\ &= \frac{6}{\sqrt{102}}\\ &\approx 0.59409\end{aligned}$$

This implies that the angle between the two planes is $\theta \approx 53.55°$. You can find the line of intersection of the two planes by simultaneously solving the two linear equations representing the planes. One way to do this is to multiply the first equation by -2 and add the result to the second equation.

$$\begin{array}{rcr} x - 2y + z = 0 & \Rightarrow & -2x + 4y - 2z = 0\\ 2x + 3y - 2z = 0 & & \underline{2x + 3y - 2z = 0}\\ & & 7y - 4z = 0 \quad \Rightarrow \quad y = \dfrac{4z}{7}\end{array}$$

Substituting $y = 4z/7$ back into one of the original equations, you can determine that $x = z/7$. Finally, by letting $t = z/7$, you obtain the parametric equations

$$x = t, \quad y = 4t, \quad \text{and} \quad z = 7t \qquad \text{Line of intersection}$$

which indicate that 1, 4, and 7 are direction numbers for the line of intersection.

Try It Exploration A

Note that the direction numbers in Example 4 can be obtained from the cross product of the two normal vectors as follows.

$$\begin{aligned}\mathbf{n}_1 \times \mathbf{n}_2 &= \begin{vmatrix}\mathbf{i} & \mathbf{j} & \mathbf{k}\\ 1 & -2 & 1\\ 2 & 3 & -2\end{vmatrix}\\ &= \begin{vmatrix}-2 & 1\\ 3 & -2\end{vmatrix}\mathbf{i} - \begin{vmatrix}1 & 1\\ 2 & -2\end{vmatrix}\mathbf{j} + \begin{vmatrix}1 & -2\\ 2 & 3\end{vmatrix}\mathbf{k}\\ &= \mathbf{i} + 4\mathbf{j} + 7\mathbf{k}\end{aligned}$$

This means that the line of intersection of the two planes is parallel to the cross product of their normal vectors.

Sketching Planes in Space

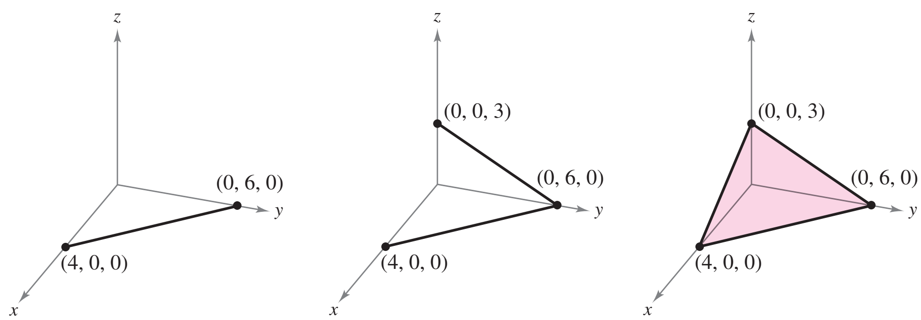
If a plane in space intersects one of the coordinate planes, the line of intersection is called the **trace** of the given plane in the coordinate plane. To sketch a plane in space, it is helpful to find its points of intersection with the coordinate axes and its traces in the coordinate planes. For example, consider the plane given by

$$3x + 2y + 4z = 12. \qquad \text{Equation of plane}$$

You can find the xy-trace by letting $z = 0$ and sketching the line

$$3x + 2y = 12 \qquad xy\text{-trace}$$

in the xy-plane. This line intersects the x-axis at $(4, 0, 0)$ and the y-axis at $(0, 6, 0)$. In Figure 11.49, this process is continued by finding the yz-trace and the xz-trace, and then shading the triangular region lying in the first octant.

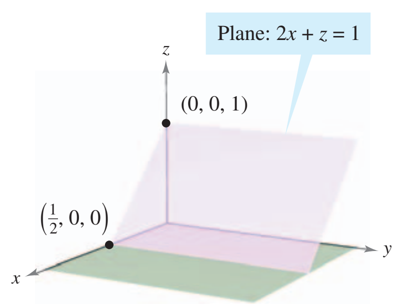


xy-trace $(z = 0)$: $3x + 2y = 12$

yz-trace $(x = 0)$: $2y + 4z = 12$

xz-trace $(y = 0)$: $3x + 4z = 12$

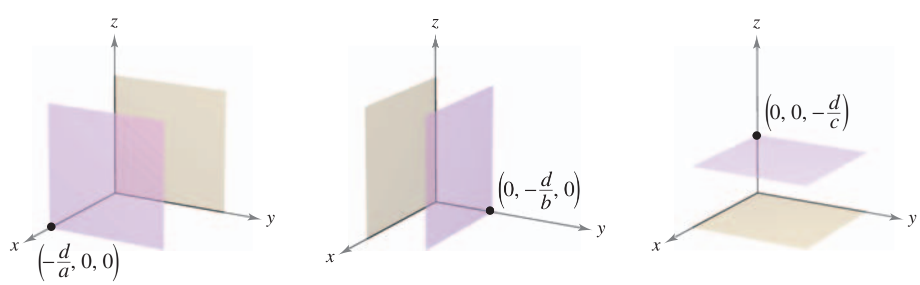
Traces of the plane $3x + 2y + 4z = 12$
Figure 11.49



Plane $2x + z = 1$ is parallel to the y-axis.
Figure 11.50

Rotatable Graph

If an equation of a plane has a missing variable, such as $2x + z = 1$, the plane must be *parallel to the axis* represented by the missing variable, as shown in Figure 11.50. If two variables are missing from an equation of a plane, it is *parallel to the coordinate plane* represented by the missing variables, as shown in Figure 11.51.



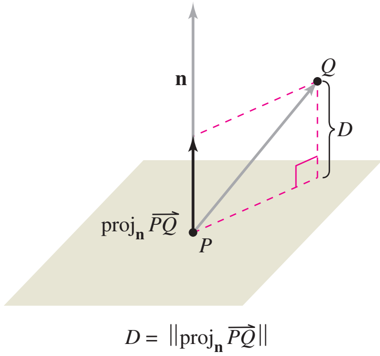
Plane $ax + d = 0$ is parallel to the yz-plane.

Plane $by + d = 0$ is parallel to the xz-plane.

Plane $cz + d = 0$ is parallel to the xy-plane.

Rotatable Graph | Rotatable Graph | Rotatable Graph

Figure 11.51



The distance between a point and a plane
Figure 11.52

Distances Between Points, Planes, and Lines

This section is concluded with the following discussion of two basic types of problems involving distance in space.

1. Finding the distance between a point and a plane
2. Finding the distance between a point and a line

The solutions of these problems illustrate the versatility and usefulness of vectors in coordinate geometry: the first problem uses the *dot product* of two vectors, and the second problem uses the *cross product*.

The distance D between a point Q and a plane is the length of the shortest line segment connecting Q to the plane, as shown in Figure 11.52. If P is *any* point in the plane, you can find this distance by projecting the vector $\overrightarrow{PQ}$ onto the normal vector $\mathbf{n}$. The length of this projection is the desired distance.

THEOREM 11.13 Distance Between a Point and a Plane

The distance between a plane and a point Q (not in the plane) is

$$D = \|\text{proj}_{\mathbf{n}}\overrightarrow{PQ}\| = \frac{|\overrightarrow{PQ} \cdot \mathbf{n}|}{\|\mathbf{n}\|}$$

where P is a point in the plane and $\mathbf{n}$ is normal to the plane.

To find a point in the plane given by $ax + by + cz + d = 0$ $(a \neq 0)$, let $y = 0$ and $z = 0$. Then, from the equation $ax + d = 0$, you can conclude that the point $(-d/a, 0, 0)$ lies in the plane.

EXAMPLE 5 Finding the Distance Between a Point and a Plane

Find the distance between the point $Q(1, 5, -4)$ and the plane given by

$$3x - y + 2z = 6.$$

Solution You know that $\mathbf{n} = \langle 3, -1, 2\rangle$ is normal to the given plane. To find a point in the plane, let $y = 0$ and $z = 0$, and obtain the point $P(2, 0, 0)$. The vector from P to Q is given by

$$\begin{aligned}\overrightarrow{PQ} &= \langle 1 - 2, 5 - 0, -4 - 0\rangle \\ &= \langle -1, 5, -4\rangle.\end{aligned}$$

Using the Distance Formula given in Theorem 11.13 produces

$$\begin{aligned}D = \frac{|\overrightarrow{PQ} \cdot \mathbf{n}|}{\|\mathbf{n}\|} &= \frac{|\langle -1, 5, -4\rangle \cdot \langle 3, -1, 2\rangle|}{\sqrt{9 + 1 + 4}} && \text{Distance between a point and a plane} \\ &= \frac{|-3 - 5 - 8|}{\sqrt{14}} \\ &= \frac{16}{\sqrt{14}}.\end{aligned}$$

Try It **Exploration A**

NOTE The choice of the point P in Example 5 is arbitrary. Try choosing a different point in the plane to verify that you obtain the same distance.

From Theorem 11.13, you can determine that the distance between the point $Q(x_0, y_0, z_0)$ and the plane given by $ax + by + cz + d = 0$ is

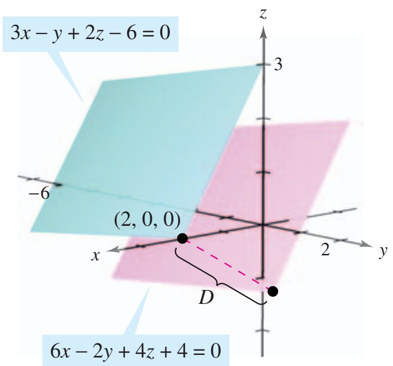
$$D = \frac{|a(x_0 - x_1) + b(y_0 - y_1) + c(z_0 - z_1)|}{\sqrt{a^2 + b^2 + c^2}}$$

or

$$D = \frac{|ax_0 + by_0 + cz_0 + d|}{\sqrt{a^2 + b^2 + c^2}}$$ Distance between a point and a plane

where $P(x_1, y_1, z_1)$ is a point in the plane and $d = -(ax_1 + by_1 + cz_1)$.

EXAMPLE 6 Finding the Distance Between Two Parallel Planes



The distance between the parallel planes is approximately 2.14.
Figure 11.53

Rotatable Graph

Find the distance between the two parallel planes given by

$$3x - y + 2z - 6 = 0 \quad \text{and} \quad 6x - 2y + 4z + 4 = 0.$$

Solution The two planes are shown in Figure 11.53. To find the distance between the planes, choose a point in the first plane, say $(x_0, y_0, z_0) = (2, 0, 0)$. Then, from the second plane, you can determine that $a = 6$, $b = -2$, $c = 4$, and $d = 4$, and conclude that the distance is

$$D = \frac{|ax_0 + by_0 + cz_0 + d|}{\sqrt{a^2 + b^2 + c^2}}$$ Distance between a point and a plane

$$= \frac{|6(2) + (-2)(0) + (4)(0) + 4|}{\sqrt{6^2 + (-2)^2 + 4^2}}$$

$$= \frac{16}{\sqrt{56}} = \frac{8}{\sqrt{14}} \approx 2.14.$$

Try It Exploration A

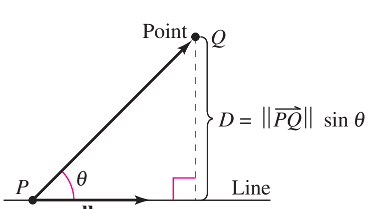
The formula for the distance between a point and a line in space resembles that for the distance between a point and a plane—except that you replace the dot product with the length of the cross product and the normal vector **n** with a direction vector for the line.

THEOREM 11.14 Distance Between a Point and a Line in Space

The distance between a point Q and a line in space is given by

$$D = \frac{\|\overrightarrow{PQ} \times \mathbf{u}\|}{\|\mathbf{u}\|}$$

where **u** is a direction vector for the line and P is a point on the line.



The distance between a point and a line
Figure 11.54

Proof In Figure 11.54, let D be the distance between the point Q and the given line. Then $D = \|\overrightarrow{PQ}\| \sin\theta$, where θ is the angle between **u** and $\overrightarrow{PQ}$. By Theorem 11.8, you have

$$\|\mathbf{u}\|\|\overrightarrow{PQ}\| \sin\theta = \|\mathbf{u} \times \overrightarrow{PQ}\| = \|\overrightarrow{PQ} \times \mathbf{u}\|.$$

Consequently,

$$D = \|\overrightarrow{PQ}\| \sin\theta = \frac{\|\overrightarrow{PQ} \times \mathbf{u}\|}{\|\mathbf{u}\|}.$$

EXAMPLE 7 Finding the Distance Between a Point and a Line

Find the distance between the point $Q(3, -1, 4)$ and the line given by

$$x = -2 + 3t, \quad y = -2t, \quad \text{and} \quad z = 1 + 4t.$$

Solution Using the direction numbers 3, -2, and 4, you know that a direction vector for the line is

$$\mathbf{u} = \langle 3, -2, 4 \rangle. \qquad \text{Direction vector for line}$$

To find a point on the line, let $t = 0$ and obtain

$$P = (-2, 0, 1). \qquad \text{Point on the line}$$

So,

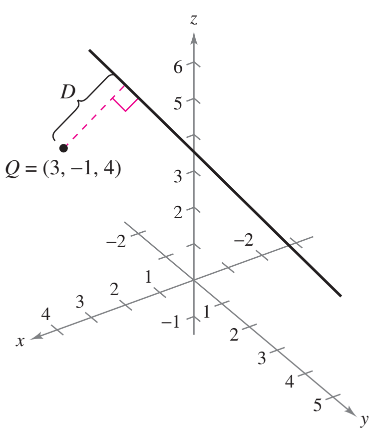
$$\overrightarrow{PQ} = \langle 3 - (-2), -1 - 0, 4 - 1 \rangle = \langle 5, -1, 3 \rangle$$

and you can form the cross product

$$\overrightarrow{PQ} \times \mathbf{u} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 5 & -1 & 3 \\ 3 & -2 & 4 \end{vmatrix} = 2\mathbf{i} - 11\mathbf{j} - 7\mathbf{k} = \langle 2, -11, -7 \rangle.$$

Finally, using Theorem 11.14, you can find the distance to be

$$D = \frac{\|\overrightarrow{PQ} \times \mathbf{u}\|}{\|\mathbf{u}\|} = \frac{\sqrt{174}}{\sqrt{29}} = \sqrt{6} \approx 2.45. \qquad \text{See Figure 11.55.}$$



The distance between the point Q and the line is $\sqrt{6} \approx 2.45$.
Figure 11.55

Try It

Exploration A

Section 11.6 Surfaces in Space

- **Recognize and write equations for cylindrical surfaces.**
- **Recognize and write equations for quadric surfaces.**
- **Recognize and write equations for surfaces of revolution.**

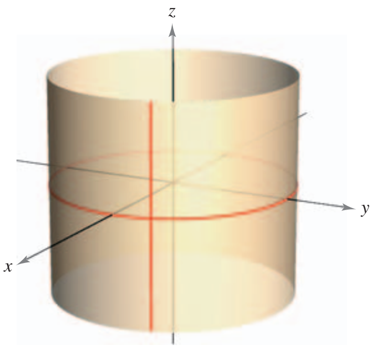
Cylindrical Surfaces

The first five sections of this chapter contained the vector portion of the preliminary work necessary to study vector calculus and the calculus of space. In this and the next section, you will study surfaces in space and alternative coordinate systems for space. You have already studied two special types of surfaces.

1. Spheres: $(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2 = r^2$ Section 11.2

2. Planes: $ax + by + cz + d = 0$ Section 11.5

A third type of surface in space is called a **cylindrical surface,** or simply a **cylinder.** To define a cylinder, consider the familiar right circular cylinder shown in Figure 11.56. You can imagine that this cylinder is generated by a vertical line moving around the circle $x^2 + y^2 = a^2$ in the xy-plane. This circle is called a **generating curve** for the cylinder, as indicated in the following definition.



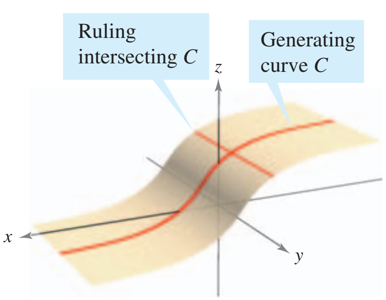
Rulings are parallel to z-axis.
Figure 11.56

Rotatable Graph

Definition of a Cylinder

Let C be a curve in a plane and let L be a line not in a parallel plane. The set of all lines parallel to L and intersecting C is called a **cylinder.** C is called the **generating curve** (or **directrix**) of the cylinder, and the parallel lines are called **rulings.**

NOTE Without loss of generality, you can assume that C lies in one of the three coordinate planes. Moreover, this text restricts the discussion to *right* cylinders—cylinders whose rulings are perpendicular to the coordinate plane containing C, as shown in Figure 11.57.



Cylinder: Rulings intersect C and are parallel to the given line.
Figure 11.57

Rotatable Graph

For the right circular cylinder shown in Figure 11.56, the equation of the generating curve is

$$x^2 + y^2 = a^2. \qquad \text{Equation of generating curve in } xy\text{-plane}$$

To find an equation for the cylinder, note that you can generate any one of the rulings by fixing the values of x and y and then allowing z to take on all real values. In this sense, the value of z is arbitrary and is, therefore, not included in the equation. In other words, the equation of this cylinder is simply the equation of its generating curve.

$$x^2 + y^2 = a^2 \qquad \text{Equation of cylinder in space}$$

Equations of Cylinders

The equation of a cylinder whose rulings are parallel to one of the coordinate axes contains only the variables corresponding to the other two axes.

EXAMPLE 1 Sketching a Cylinder

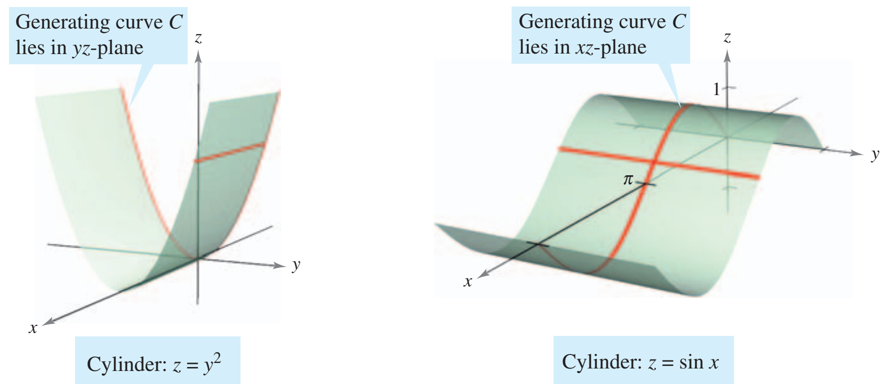
Sketch the surface represented by each equation.

a. $z = y^2$ **b.** $z = \sin x, \quad 0 \le x \le 2\pi$

Solution

a. The graph is a cylinder whose generating curve, $z = y^2$, is a parabola in the yz-plane. The rulings of the cylinder are parallel to the x-axis, as shown in Figure 11.58(a).

b. The graph is a cylinder generated by the sine curve in the xz-plane. The rulings are parallel to the y-axis, as shown in Figure 11.58(b).



(a) Rulings are parallel to x-axis. **(b)** Rulings are parallel to y-axis.

Rotatable Graph Rotatable Graph

Figure 11.58

Try It Exploration A

Quadric Surfaces

The fourth basic type of surface in space is a **quadric surface.** Quadric surfaces are the three-dimensional analogs of conic sections.

STUDY TIP In the table on pages 812 and 813, only one of several orientations of each quadric surface is shown. If the surface is oriented along a different axis, its standard equation will change accordingly, as illustrated in Examples 2 and 3. The fact that the two types of paraboloids have one variable raised to the first power can be helpful in classifying quadric surfaces. The other four types of basic quadric surfaces have equations that are of *second degree* in all three variables.

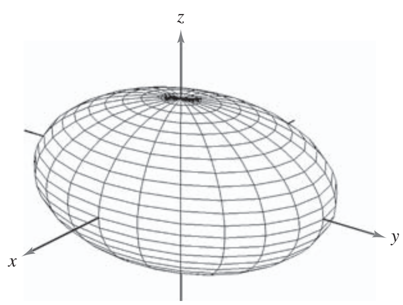
Quadric Surface

The equation of a **quadric surface** in space is a second-degree equation of the form

$$Ax^2 + By^2 + Cz^2 + Dxy + Exz + Fyz + Gx + Hy + Iz + J = 0.$$

There are six basic types of quadric surfaces: **ellipsoid, hyperboloid of one sheet, hyperboloid of two sheets, elliptic cone, elliptic paraboloid, and hyperbolic paraboloid.**

The intersection of a surface with a plane is called the **trace of the surface** in the plane. To visualize a surface in space, it is helpful to determine its traces in some well-chosen planes. The traces of quadric surfaces are conics. These traces, together with the **standard form** of the equation of each quadric surface, are shown in the table on pages 812 and 813.



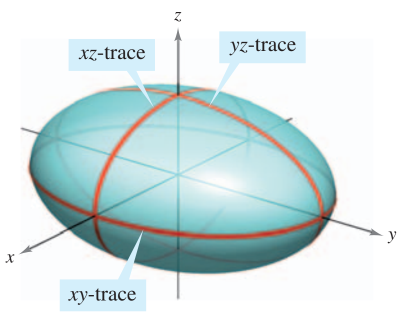
Rotatable Graph

Ellipsoid

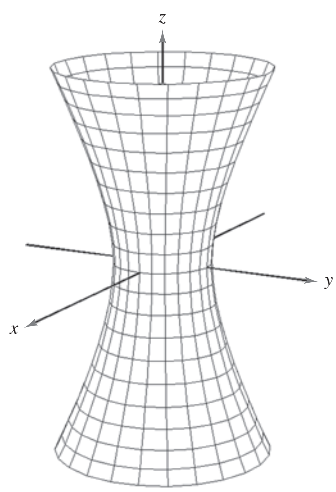
$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$$

Trace	*Plane*
Ellipse	Parallel to xy-plane
Ellipse	Parallel to xz-plane
Ellipse	Parallel to yz-plane

The surface is a sphere if $a = b = c \neq 0$.



Rotatable Graph



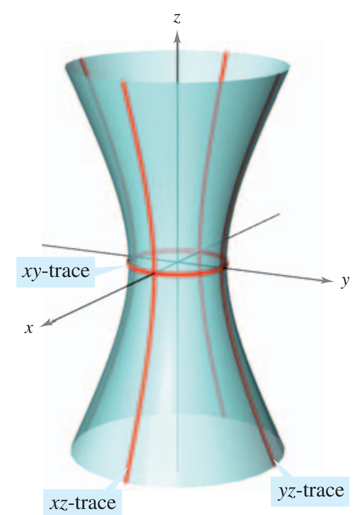
Rotatable Graph

Hyperboloid of One Sheet

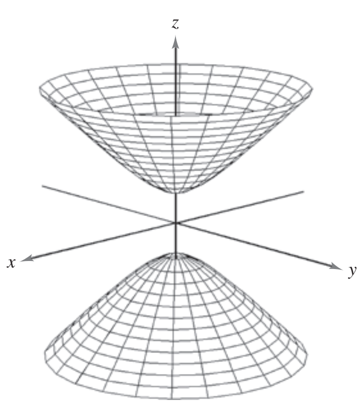
$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1$$

Trace	*Plane*
Ellipse	Parallel to xy-plane
Hyperbola	Parallel to xz-plane
Hyperbola	Parallel to yz-plane

The axis of the hyperboloid corresponds to the variable whose coefficient is negative.



Rotatable Graph



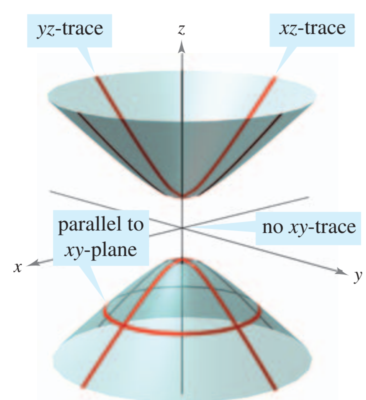
Rotatable Graph

Hyperboloid of Two Sheets

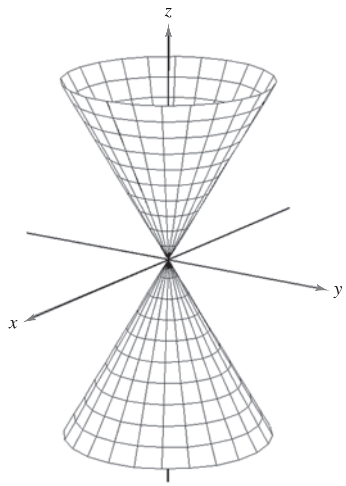
$$\frac{z^2}{c^2} - \frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$$

Trace	*Plane*
Ellipse	Parallel to xy-plane
Hyperbola	Parallel to xz-plane
Hyperbola	Parallel to yz-plane

The axis of the hyperboloid corresponds to the variable whose coefficient is positive. There is no trace in the coordinate plane perpendicular to this axis.



Rotatable Graph



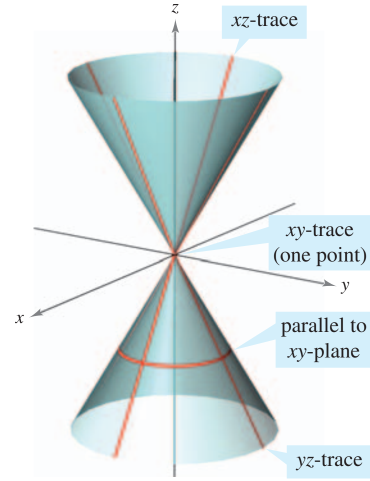
Rotatable Graph

Elliptic Cone

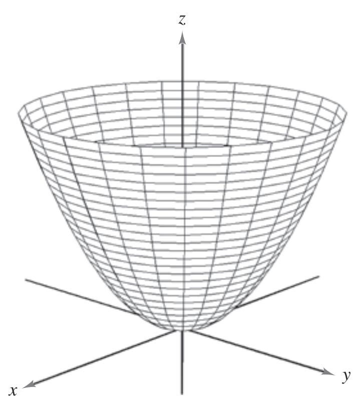
$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 0$$

Trace	*Plane*
Ellipse	Parallel to xy-plane
Hyperbola	Parallel to xz-plane
Hyperbola	Parallel to yz-plane

The axis of the cone corresponds to the variable whose coefficient is negative. The traces in the coordinate planes parallel to this axis are intersecting lines.



Rotatable Graph



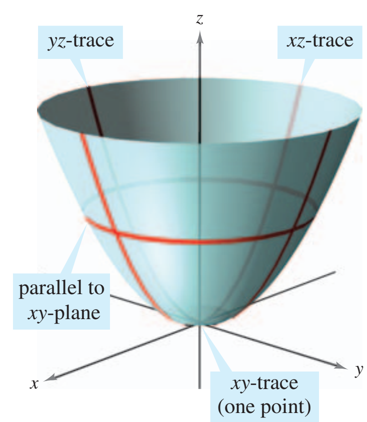
Rotatable Graph

Elliptic Paraboloid

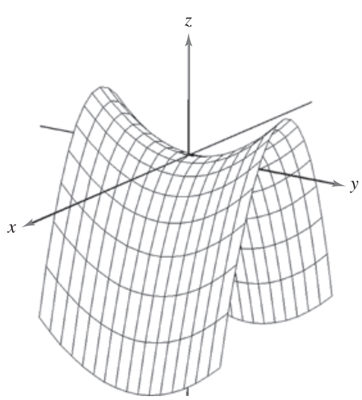
$$z = \frac{x^2}{a^2} + \frac{y^2}{b^2}$$

Trace	*Plane*
Ellipse	Parallel to xy-plane
Parabola	Parallel to xz-plane
Parabola	Parallel to yz-plane

The axis of the paraboloid corresponds to the variable raised to the first power.



Rotatable Graph



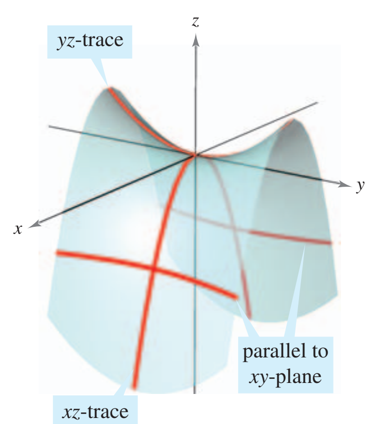
Rotatable Graph

Hyperbolic Paraboloid

$$z = \frac{y^2}{b^2} - \frac{x^2}{a^2}$$

Trace	*Plane*
Hyperbola	Parallel to xy-plane
Parabola	Parallel to xz-plane
Parabola	Parallel to yz-plane

The axis of the paraboloid corresponds to the variable raised to the first power.



Rotatable Graph

To classify a quadric surface, begin by writing the surface in standard form. Then, determine several traces taken in the coordinate planes *or* taken in planes that are parallel to the coordinate planes.

EXAMPLE 2 Sketching a Quadric Surface

Classify and sketch the surface given by $4x^2 - 3y^2 + 12z^2 + 12 = 0$.

Solution Begin by writing the equation in standard form.

$$4x^2 - 3y^2 + 12z^2 + 12 = 0 \qquad \text{Write original equation.}$$

$$\frac{x^2}{-3} + \frac{y^2}{4} - z^2 - 1 = 0 \qquad \text{Divide by } -12.$$

$$\frac{y^2}{4} - \frac{x^2}{3} - \frac{z^2}{1} = 1 \qquad \text{Standard form}$$

From the table on pages 812 and 813, you can conclude that the surface is a hyperboloid of two sheets with the y-axis as its axis. To sketch the graph of this surface, it helps to find the traces in the coordinate planes.

$$xy\text{-trace } (z = 0): \quad \frac{y^2}{4} - \frac{x^2}{3} = 1 \qquad \text{Hyperbola}$$

$$xz\text{-trace } (y = 0): \quad \frac{x^2}{3} + \frac{z^2}{1} = -1 \qquad \text{No trace}$$

$$yz\text{-trace } (x = 0): \quad \frac{y^2}{4} - \frac{z^2}{1} = 1 \qquad \text{Hyperbola}$$

The graph is shown in Figure 11.59.

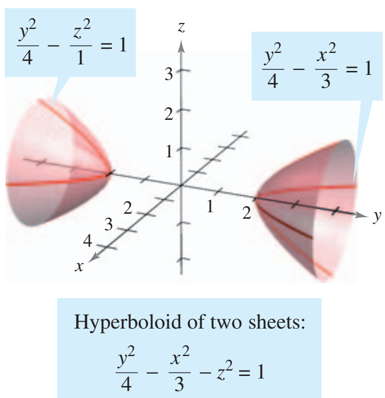


Figure 11.59

Rotatable Graph

Try It Exploration A

EXAMPLE 3 Sketching a Quadric Surface

Classify and sketch the surface given by $x - y^2 - 4z^2 = 0$.

Solution Because x is raised only to the first power, the surface is a paraboloid. The axis of the paraboloid is the x-axis. In the standard form, the equation is

$$x = y^2 + 4z^2. \qquad \text{Standard form}$$

Some convenient traces are as follows.

$$xy\text{-trace } (z = 0): \quad x = y^2 \qquad \text{Parabola}$$

$$xz\text{-trace } (y = 0): \quad x = 4z^2 \qquad \text{Parabola}$$

$$\text{parallel to } yz\text{-plane } (x = 4): \quad \frac{y^2}{4} + \frac{z^2}{1} = 1 \qquad \text{Ellipse}$$

The surface is an *elliptic* paraboloid, as shown in Figure 11.60.

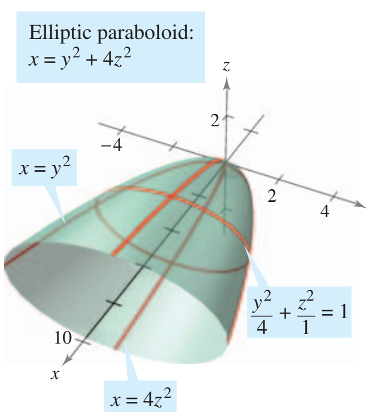


Figure 11.60

Rotatable Graph

Try It Exploration A

Some second-degree equations in x, y, and z do not represent any of the basic types of quadric surfaces. Here are two examples.

$$x^2 + y^2 + z^2 = 0 \qquad \text{Single point}$$

$$x^2 + y^2 = 1 \qquad \text{Right circular cylinder}$$

For a quadric surface not centered at the origin, you can form the standard equation by completing the square, as demonstrated in Example 4.

EXAMPLE 4 A Quadric Surface Not Centered at the Origin

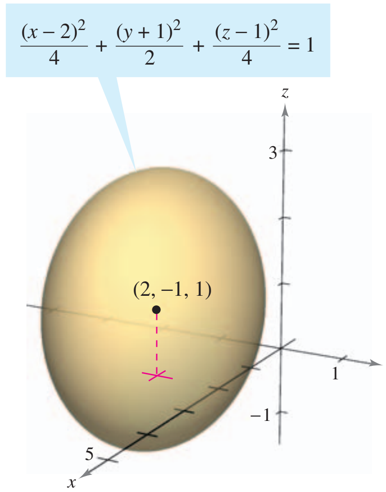
Classify and sketch the surface given by

$$x^2 + 2y^2 + z^2 - 4x + 4y - 2z + 3 = 0.$$

Solution Completing the square for each variable produces the following.

$$(x^2 - 4x + \quad) + 2(y^2 + 2y + \quad) + (z^2 - 2z + \quad) = -3$$
$$(x^2 - 4x + 4) + 2(y^2 + 2y + 1) + (z^2 - 2z + 1) = -3 + 4 + 2 + 1$$
$$(x - 2)^2 + 2(y + 1)^2 + (z - 1)^2 = 4$$
$$\frac{(x-2)^2}{4} + \frac{(y+1)^2}{2} + \frac{(z-1)^2}{4} = 1$$

From this equation, you can see that the quadric surface is an ellipsoid that is centered at $(2, -1, 1)$. Its graph is shown in Figure 11.61.



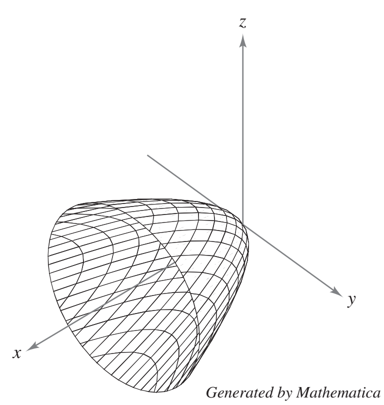
An ellipsoid centered at $(2, -1, 1)$
Figure 11.61

Rotatable Graph

Try It | Exploration A | Open Exploration

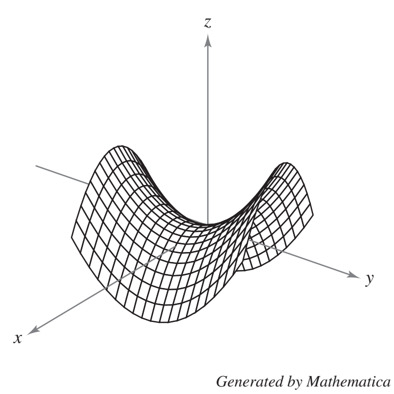
TECHNOLOGY

A computer algebra system can help you visualize a surface in space.* Most of these computer algebra systems create three-dimensional illusions by sketching several traces of the surface and then applying a "hidden-line" routine that blocks out portions of the surface that lie behind other portions of the surface. Two examples of figures that were generated by *Mathematica* are shown below.



Rotatable Graph

Elliptic paraboloid
$x = \frac{y^2}{2} + \frac{z^2}{2}$



Rotatable Graph

Hyperbolic paraboloid
$z = \frac{y^2}{16} - \frac{x^2}{16}$

Using a graphing utility to graph a surface in space requires practice. For one thing, you must know enough about the surface to be able to specify a *viewing window* that gives a representative view of the surface. Also, you can often improve the view of a surface by rotating the axes. For instance, note that the elliptic paraboloid in the figure is seen from a line of sight that is "higher" than the line of sight used to view the hyperbolic paraboloid.

**Some 3-D graphing utilities require surfaces to be entered with parametric equations. For a discussion of this technique, see Section 15.5.*

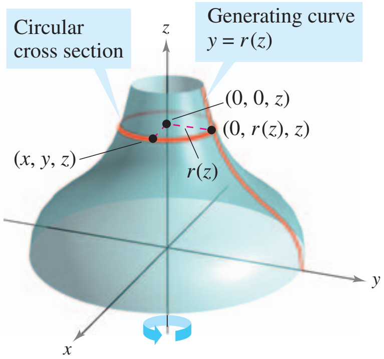


Figure 11.62

Rotatable Graph

Surfaces of Revolution

The fifth special type of surface you will study is called a **surface of revolution.** In Section 7.4, you studied a method for finding the *area* of such a surface. You will now look at a procedure for finding its *equation*. Consider the graph of the **radius function**

$$y = r(z) \qquad \text{Generating curve}$$

in the yz-plane. If this graph is revolved about the z-axis, it forms a surface of revolution, as shown in Figure 11.62. The trace of the surface in the plane $z = z_0$ is a circle whose radius is $r(z_0)$ and whose equation is

$$x^2 + y^2 = [r(z_0)]^2. \qquad \text{Circular trace in plane: } z = z_0$$

Replacing z_0 with z produces an equation that is valid for all values of z. In a similar manner, you can obtain equations for surfaces of revolution for the other two axes, and the results are summarized as follows.

Surface of Revolution

If the graph of a radius function r is revolved about one of the coordinate axes, the equation of the resulting surface of revolution has one of the following forms.

1. Revolved about the x-axis: $y^2 + z^2 = [r(x)]^2$
2. Revolved about the y-axis: $x^2 + z^2 = [r(y)]^2$
3. Revolved about the z-axis: $x^2 + y^2 = [r(z)]^2$

EXAMPLE 5 Finding an Equation for a Surface of Revolution

a. An equation for the surface of revolution formed by revolving the graph of

$$y = \frac{1}{z} \qquad \text{Radius function}$$

about the z-axis is

$$x^2 + y^2 = [r(z)]^2 \qquad \text{Revolved about the } z\text{-axis}$$

$$x^2 + y^2 = \left(\frac{1}{z}\right)^2. \qquad \text{Substitute } 1/z \text{ for } r(z).$$

b. To find an equation for the surface formed by revolving the graph of $9x^2 = y^3$ about the y-axis, solve for x in terms of y to obtain

$$x = \tfrac{1}{3}y^{3/2} = r(y). \qquad \text{Radius function}$$

So, the equation for this surface is

$$x^2 + z^2 = [r(y)]^2 \qquad \text{Revolved about the } y\text{-axis}$$

$$x^2 + z^2 = \left(\tfrac{1}{3}y^{3/2}\right)^2 \qquad \text{Substitute } \tfrac{1}{3}y^{3/2} \text{ for } r(y).$$

$$x^2 + z^2 = \tfrac{1}{9}y^3. \qquad \text{Equation of surface}$$

The graph is shown in Figure 11.63.

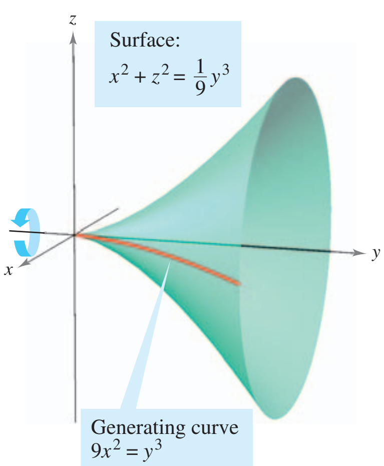


Figure 11.63

Rotatable Graph

Try It

Exploration A

The generating curve for a surface of revolution is not unique. For instance, the surface

$$x^2 + z^2 = e^{-2y}$$

can be formed by revolving either the graph of $x = e^{-y}$ about the y-axis or the graph of $z = e^{-y}$ about the y-axis, as shown in Figure 11.64.

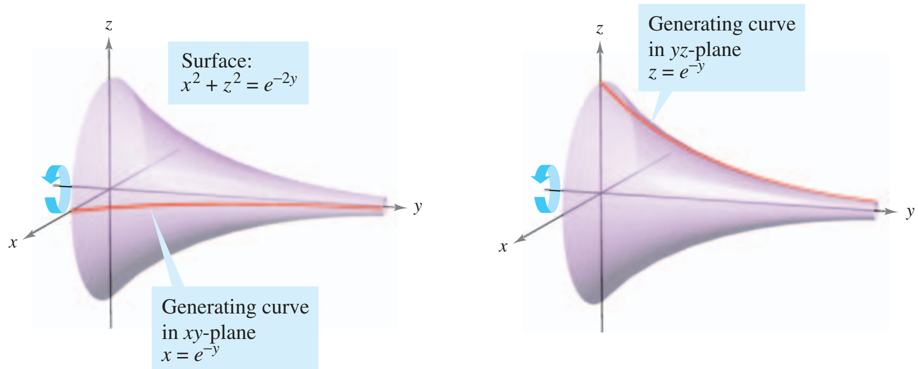


Figure 11.64

Rotatable Graph

EXAMPLE 6 Finding a Generating Curve for a Surface of Revolution

Find a generating curve and the axis of revolution for the surface given by

$$x^2 + 3y^2 + z^2 = 9.$$

Solution You now know that the equation has one of the following forms.

$$x^2 + y^2 = [r(z)]^2 \qquad \text{Revolved about } z\text{-axis}$$

$$y^2 + z^2 = [r(x)]^2 \qquad \text{Revolved about } x\text{-axis}$$

$$x^2 + z^2 = [r(y)]^2 \qquad \text{Revolved about } y\text{-axis}$$

Because the coefficients of x^2 and z^2 are equal, you should choose the third form and write

$$x^2 + z^2 = 9 - 3y^2.$$

The y-axis is the axis of revolution. You can choose a generating curve from either of the following traces.

$$x^2 = 9 - 3y^2 \qquad \text{Trace in } xy\text{-plane}$$

$$z^2 = 9 - 3y^2 \qquad \text{Trace in } yz\text{-plane}$$

For example, using the first trace, the generating curve is the semiellipse given by

$$x = \sqrt{9 - 3y^2}. \qquad \text{Generating curve}$$

The graph of this surface is shown in Figure 11.65.

Try It Exploration A

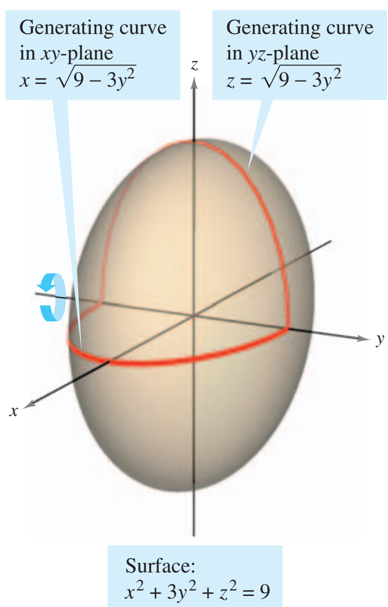


Figure 11.65

Rotatable Graph

Section 11.7 Cylindrical and Spherical Coordinates

- **Use cylindrical coordinates to represent surfaces in space.**
- **Use spherical coordinates to represent surfaces in space.**

Cylindrical Coordinates

You have already seen that some two-dimensional graphs are easier to represent in polar coordinates than in rectangular coordinates. A similar situation exists for surfaces in space. In this section, you will study two alternative space-coordinate systems. The first, the **cylindrical coordinate system,** is an extension of polar coordinates in the plane to three-dimensional space.

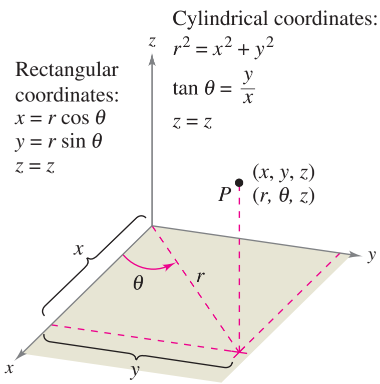


Figure 11.66

> **The Cylindrical Coordinate System**
>
> In a **cylindrical coordinate system,** a point P in space is represented by an ordered triple (r, θ, z).
>
> **1.** (r, θ) is a polar representation of the projection of P in the xy-plane.
>
> **2.** z is the directed distance from (r, θ) to P.

To convert from rectangular to cylindrical coordinates (or vice versa), use the following conversion guidelines for polar coordinates, as illustrated in Figure 11.66.

Cylindrical to rectangular:

$$x = r\cos\theta, \qquad y = r\sin\theta, \qquad z = z$$

Rectangular to cylindrical:

$$r^2 = x^2 + y^2, \qquad \tan\theta = \frac{y}{x}, \qquad z = z$$

The point $(0, 0, 0)$ is called the **pole.** Moreover, because the representation of a point in the polar coordinate system is not unique, it follows that the representation in the cylindrical coordinate system is also not unique.

EXAMPLE 1 Converting from Cylindrical to Rectangular Coordinates

Convert the point $(r, \theta, z) = \left(4, \frac{5\pi}{6}, 3\right)$ to rectangular coordinates.

Solution Using the cylindrical-to-rectangular conversion equations produces

$$x = 4\cos\frac{5\pi}{6} = 4\left(-\frac{\sqrt{3}}{2}\right) = -2\sqrt{3}$$

$$y = 4\sin\frac{5\pi}{6} = 4\left(\frac{1}{2}\right) = 2$$

$$z = 3.$$

So, in rectangular coordinates, the point is $(x, y, z) = \left(-2\sqrt{3}, 2, 3\right)$, as shown in Figure 11.67.

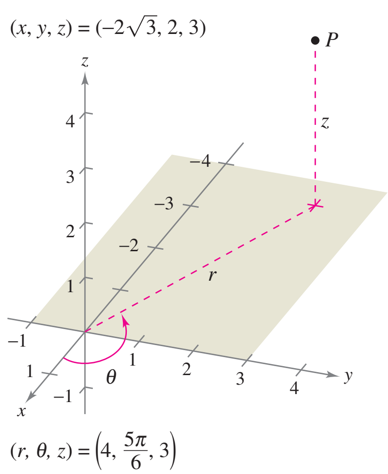


Figure 11.67

Try It

Exploration A

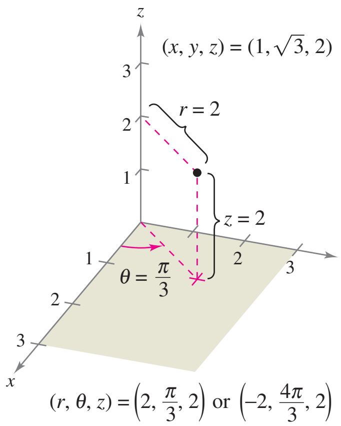


Figure 11.68

EXAMPLE 2 Converting from Rectangular to Cylindrical Coordinates

Convert the point $(x, y, z) = \left(1, \sqrt{3}, 2\right)$ to cylindrical coordinates.

Solution Use the rectangular-to-cylindrical conversion equations.

$$r = \pm\sqrt{1 + 3} = \pm 2$$

$$\tan\theta = \sqrt{3} \quad \Rightarrow \quad \theta = \arctan\left(\sqrt{3}\right) + n\pi = \frac{\pi}{3} + n\pi$$

$$z = 2$$

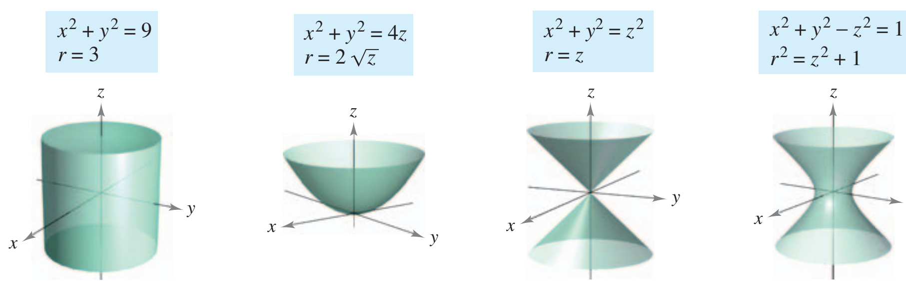
You have two choices for r and infinitely many choices for θ. As shown in Figure 11.68, two convenient representations of the point are

$$\left(2, \frac{\pi}{3}, 2\right) \qquad r > 0 \text{ and } \theta \text{ in Quadrant I}$$

$$\left(-2, \frac{4\pi}{3}, 2\right). \qquad r < 0 \text{ and } \theta \text{ in Quadrant III}$$

Try It   Exploration A

Cylindrical coordinates are especially convenient for representing cylindrical surfaces and surfaces of revolution with the z-axis as the axis of symmetry, as shown in Figure 11.69.

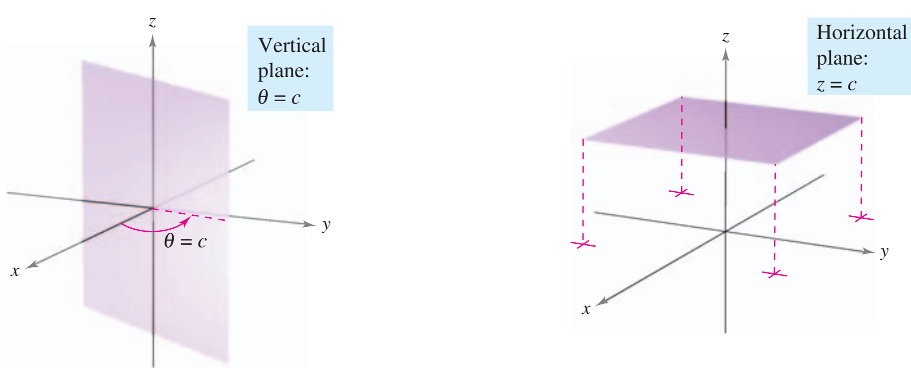


Cylinder   Paraboloid   Cone   Hyperboloid

Rotatable Graph   Rotatable Graph   Rotatable Graph   Rotatable Graph

Figure 11.69

Vertical planes containing the z-axis and horizontal planes also have simple cylindrical coordinate equations, as shown in Figure 11.70.



Rotatable Graph   Rotatable Graph

Figure 11.70

EXAMPLE 3 Rectangular-to-Cylindrical Conversion

Find an equation in cylindrical coordinates for the surface represented by each rectangular equation.

a. $x^2 + y^2 = 4z^2$

b. $y^2 = x$

Solution

a. From the preceding section, you know that the graph $x^2 + y^2 = 4z^2$ is a "double-napped" cone with its axis along the z-axis, as shown in Figure 11.71. If you replace $x^2 + y^2$ with r^2, the equation in cylindrical coordinates is

$$x^2 + y^2 = 4z^2 \qquad \text{Rectangular equation}$$
$$r^2 = 4z^2. \qquad \text{Cylindrical equation}$$

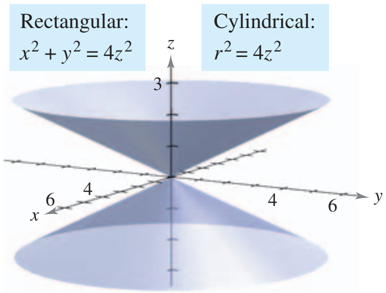


Figure 11.71

Rotatable Graph

b. The graph of the surface $y^2 = x$ is a parabolic cylinder with rulings parallel to the z-axis, as shown in Figure 11.72. By replacing y^2 with $r^2 \sin^2 \theta$ and x with $r \cos \theta$, you obtain the following equation in cylindrical coordinates.

$$y^2 = x \qquad \text{Rectangular equation}$$
$$r^2 \sin^2 \theta = r \cos \theta \qquad \text{Substitute } r \sin \theta \text{ for } y \text{ and } r \cos \theta \text{ for } x.$$
$$r(r \sin^2 \theta - \cos \theta) = 0 \qquad \text{Collect terms and factor.}$$
$$r \sin^2 \theta - \cos \theta = 0 \qquad \text{Divide each side by } r.$$
$$r = \frac{\cos \theta}{\sin^2 \theta} \qquad \text{Solve for } r.$$
$$r = \csc \theta \cot \theta \qquad \text{Cylindrical equation}$$

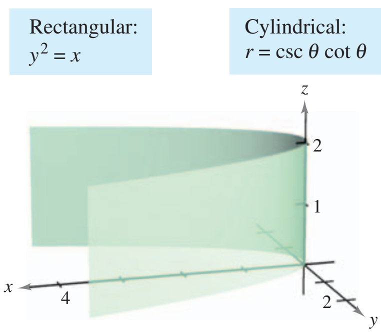


Figure 11.72

Rotatable Graph

Note that this equation includes a point for which $r = 0$, so nothing was lost by dividing each side by the factor r.

Try It Exploration A

Converting from rectangular coordinates to cylindrical coordinates is more straightforward than converting from cylindrical coordinates to rectangular coordinates, as demonstrated in Example 4.

EXAMPLE 4 Cylindrical-to-Rectangular Conversion

Find an equation in rectangular coordinates for the surface represented by the cylindrical equation

$$r^2 \cos 2\theta + z^2 + 1 = 0.$$

Solution

$$r^2 \cos 2\theta + z^2 + 1 = 0 \qquad \text{Cylindrical equation}$$
$$r^2(\cos^2 \theta - \sin^2 \theta) + z^2 + 1 = 0 \qquad \text{Trigonometric identity}$$
$$r^2 \cos^2 \theta - r^2 \sin^2 \theta + z^2 = -1$$
$$x^2 - y^2 + z^2 = -1 \qquad \text{Replace } r \cos \theta \text{ with } x \text{ and } r \sin \theta \text{ with } y.$$
$$y^2 - x^2 - z^2 = 1 \qquad \text{Rectangular equation}$$

This is a hyperboloid of two sheets whose axis lies along the y-axis, as shown in Figure 11.73.

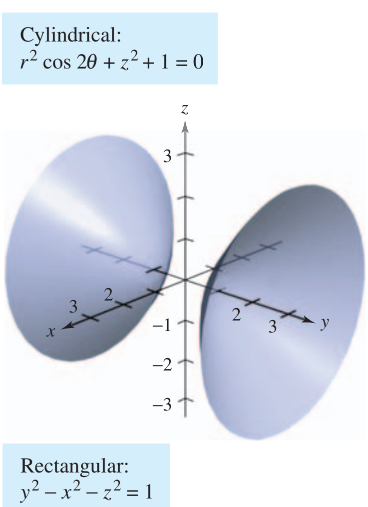


Figure 11.73

Rotatable Graph

Try It Exploration A

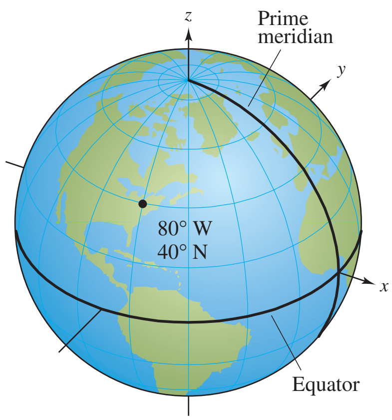


Figure 11.74

Spherical Coordinates

In the **spherical coordinate system,** each point is represented by an ordered triple: the first coordinate is a distance, and the second and third coordinates are angles. This system is similar to the latitude-longitude system used to identify points on the surface of Earth. For example, the point on the surface of Earth whose latitude is 40° North (of the equator) and whose longitude is 80° West (of the prime meridian) is shown in Figure 11.74. Assuming that the Earth is spherical and has a radius of 4000 miles, you would label this point as

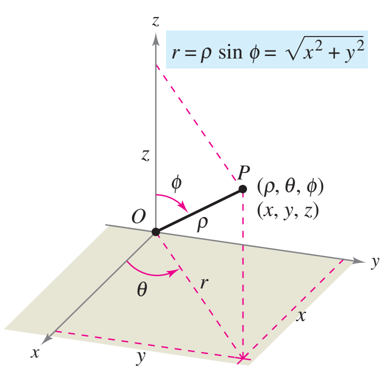
$$\underbrace{(4000}_{\text{Radius}},\ \underbrace{-80°}_{\text{80° clockwise from prime meridian}},\ \underbrace{50°)}_{\text{50° down from North Pole}}.$$

The Spherical Coordinate System

In a **spherical coordinate system,** a point P in space is represented by an ordered triple (ρ, θ, ϕ).

1. ρ is the distance between P and the origin, $\rho \geq 0$.
2. θ is the same angle used in cylindrical coordinates for $r \geq 0$.
3. ϕ is the angle *between* the positive z-axis and the line segment $\overrightarrow{OP}$, $0 \leq \phi \leq \pi$.

Note that the first and third coordinates, ρ and ϕ, are nonnegative. ρ is the lowercase Greek letter *rho*, and ϕ is the lowercase Greek letter *phi*.



Spherical coordinates
Figure 11.75

The relationship between rectangular and spherical coordinates is illustrated in Figure 11.75. To convert from one system to the other, use the following.

Spherical to rectangular:

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

Rectangular to spherical:

$$\rho^2 = x^2 + y^2 + z^2, \qquad \tan\theta = \frac{y}{x}, \qquad \phi = \arccos\left(\frac{z}{\sqrt{x^2 + y^2 + z^2}}\right)$$

To change coordinates between the cylindrical and spherical systems, use the following.

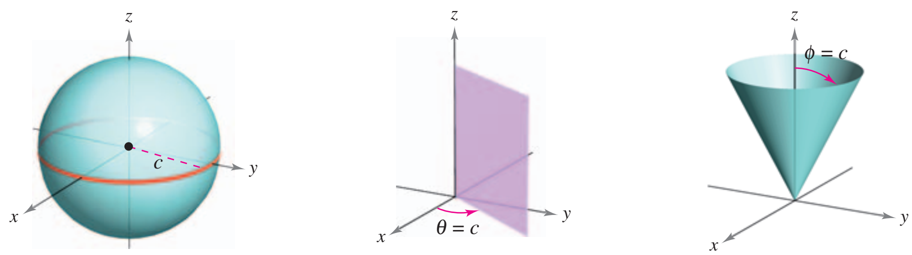
Spherical to cylindrical (r ≥ 0):

$$r^2 = \rho^2 \sin^2\phi, \qquad \theta = \theta, \qquad z = \rho\cos\phi$$

Cylindrical to spherical (r ≥ 0):

$$\rho = \sqrt{r^2 + z^2}, \qquad \theta = \theta, \qquad \phi = \arccos\left(\frac{z}{\sqrt{r^2 + z^2}}\right)$$

The spherical coordinate system is useful primarily for surfaces in space that have a *point* or *center* of symmetry. For example, Figure 11.76 shows three surfaces with simple spherical equations.



Sphere: $\rho = c$

Vertical half-plane: $\theta = c$

Half-cone: $\phi = c$ $\left(0 < c < \frac{\pi}{2}\right)$

Rotatable Graph | Rotatable Graph | Rotatable Graph

Figure 11.76

EXAMPLE 5 Rectangular-to-Spherical Conversion

Find an equation in spherical coordinates for the surface represented by each rectangular equation.

a. Cone: $x^2 + y^2 = z^2$

b. Sphere: $x^2 + y^2 + z^2 - 4z = 0$

Solution

a. Making the appropriate replacements for x, y, and z in the given equation yields the following.

$$
\begin{aligned}
x^2 + y^2 &= z^2 \\
\rho^2 \sin^2 \phi \cos^2 \theta + \rho^2 \sin^2 \phi \sin^2 \theta &= \rho^2 \cos^2 \phi \\
\rho^2 \sin^2 \phi (\cos^2 \theta + \sin^2 \theta) &= \rho^2 \cos^2 \phi \\
\rho^2 \sin^2 \phi &= \rho^2 \cos^2 \phi \\
\frac{\sin^2 \phi}{\cos^2 \phi} &= 1 && \rho \geq 0 \\
\tan^2 \phi &= 1 && \phi = \pi/4 \text{ or } \phi = 3\pi/4
\end{aligned}
$$

The equation $\phi = \pi/4$ represents the *upper* half-cone, and the equation $\phi = 3\pi/4$ represents the *lower* half-cone.

b. Because $\rho^2 = x^2 + y^2 + z^2$ and $z = \rho \cos \phi$, the given equation has the following spherical form.

$$\rho^2 - 4\rho \cos \phi = 0 \quad \Rightarrow \quad \rho(\rho - 4 \cos \phi) = 0$$

Temporarily discarding the possibility that $\rho = 0$, you have the spherical equation

$$\rho - 4 \cos \phi = 0 \qquad \text{or} \qquad \rho = 4 \cos \phi.$$

Note that the solution set for this equation includes a point for which $\rho = 0$, so nothing is lost by discarding the factor ρ. The sphere represented by the equation $\rho = 4 \cos \phi$ is shown in Figure 11.77.

Rectangular: $x^2 + y^2 + z^2 - 4z = 0$

Spherical: $\rho = 4 \cos \phi$

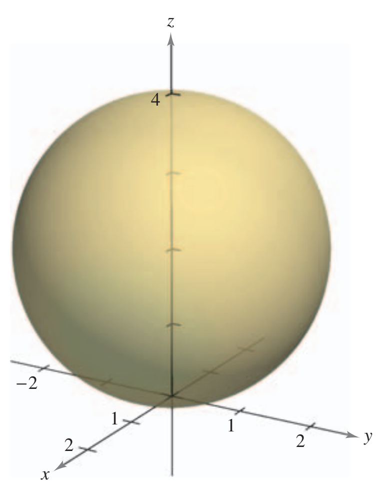


Figure 11.77

Rotatable Graph

Try It | Exploration A | Exploration B | Open Exploration

Section 12.1 Vector-Valued Functions

- **Analyze and sketch a space curve given by a vector-valued function.**
- **Extend the concepts of limits and continuity to vector-valued functions.**

Space Curves and Vector-Valued Functions

In Section 10.2, a *plane curve* was defined as the set of ordered pairs $(f(t), g(t))$ together with their defining parametric equations

$$x = f(t) \qquad \text{and} \qquad y = g(t)$$

where f and g are continuous functions of t on an interval I. This definition can be extended naturally to three-dimensional space as follows. A **space curve** C is the set of all ordered triples $(f(t), g(t), h(t))$ together with their defining parametric equations

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

where f, g, and h are continuous functions of t on an interval I.

Before looking at examples of space curves, a new type of function, called a **vector-valued function,** is introduced. This type of function maps real numbers to vectors.

Definition of Vector-Valued Function

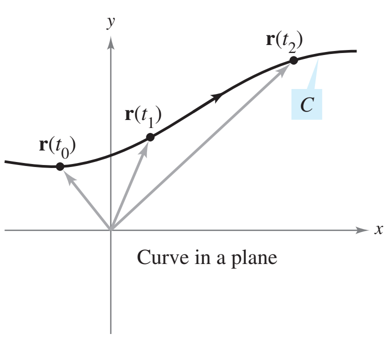
A function of the form

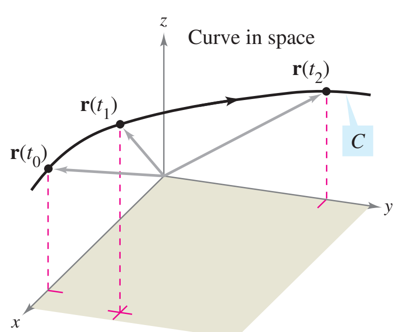
$$\mathbf{r}(t) = f(t)\mathbf{i} + g(t)\mathbf{j} \qquad \text{Plane}$$

or

$$\mathbf{r}(t) = f(t)\mathbf{i} + g(t)\mathbf{j} + h(t)\mathbf{k} \qquad \text{Space}$$

is a **vector-valued function,** where the **component functions** f, g, and h are real-valued functions of the parameter t. Vector-valued functions are sometimes denoted as $\mathbf{r}(t) = \langle f(t), g(t)\rangle$ or $\mathbf{r}(t) = \langle f(t), g(t), h(t)\rangle$.





Curve C is traced out by the terminal point of position vector $\mathbf{r}(t)$.
Figure 12.1

Technically, a curve in the plane or in space consists of a collection of points and the defining parametric equations. Two different curves can have the same graph. For instance, each of the curves given by

$$\mathbf{r} = \sin t\,\mathbf{i} + \cos t\,\mathbf{j} \qquad \text{and} \qquad \mathbf{r} = \sin t^2\,\mathbf{i} + \cos t^2\,\mathbf{j}$$

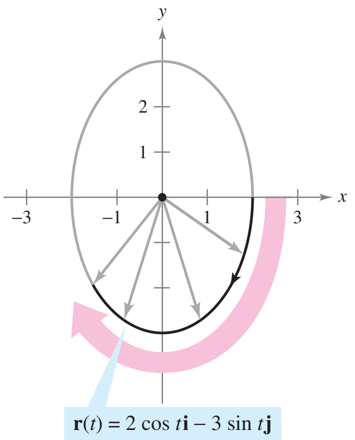
has the unit circle as its graph, but these equations do not represent the same curve—because the circle is traced out in different ways on the graphs. To see that these equations do not represent the same curve, select the animation button.

Animation

Be sure you see the distinction between the vector-valued function $\mathbf{r}$ and the real-valued functions f, g, and h. All are functions of the real variable t, but $\mathbf{r}(t)$ is a vector, whereas $f(t)$, $g(t)$, and $h(t)$ are real numbers (for each specific value of t).

Vector-valued functions serve dual roles in the representation of curves. By letting the parameter t represent time, you can use a vector-valued function to represent *motion* along a curve. Or, in the more general case, you can use a vector-valued function to *trace the graph* of a curve. In either case, the terminal point of the position vector $\mathbf{r}(t)$ coincides with the point (x, y) or (x, y, z) on the curve given by the parametric equations, as shown in Figure 12.1. The arrowhead on the curve indicates the curve's *orientation* by pointing in the direction of increasing values of t.

Unless stated otherwise, the **domain** of a vector-valued function $\mathbf{r}$ is considered to be the intersection of the domains of the component functions f, g, and h. For instance, the domain of $\mathbf{r}(t) = (\ln t)\mathbf{i} + \sqrt{1-t}\,\mathbf{j} + t\mathbf{k}$ is the interval $(0, 1]$.



The ellipse is traced clockwise as t increases from 0 to 2π.
Figure 12.2

Animation

EXAMPLE 1 Sketching a Plane Curve

Sketch the plane curve represented by the vector-valued function

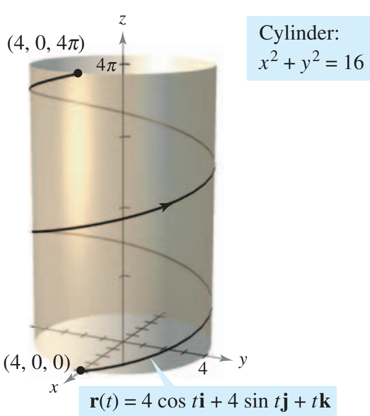
$$\mathbf{r}(t) = 2\cos t\,\mathbf{i} - 3\sin t\,\mathbf{j}, \quad 0 \le t \le 2\pi. \qquad \text{Vector-valued function}$$

Solution From the position vector $\mathbf{r}(t)$, you can write the parametric equations $x = 2\cos t$ and $y = -3\sin t$. Solving for $\cos t$ and $\sin t$ and using the identity $\cos^2 t + \sin^2 t = 1$ produces the rectangular equation

$$\frac{x^2}{2^2} + \frac{y^2}{3^2} = 1. \qquad \text{Rectangular equation}$$

The graph of this rectangular equation is the ellipse shown in Figure 12.2. The curve has a *clockwise* orientation. That is, as t increases from 0 to 2π, the position vector $\mathbf{r}(t)$ moves clockwise, and its terminal point traces the ellipse.

Try It | Exploration A | Exploration B



As t increases from 0 to 4π, two spirals on the helix are traced out.
Figure 12.3

Rotatable Graph

EXAMPLE 2 Sketching a Space Curve

Sketch the space curve represented by the vector-valued function

$$\mathbf{r}(t) = 4\cos t\,\mathbf{i} + 4\sin t\,\mathbf{j} + t\mathbf{k}, \quad 0 \le t \le 4\pi. \qquad \text{Vector-valued function}$$

Solution From the first two parametric equations $x = 4\cos t$ and $y = 4\sin t$, you can obtain

$$x^2 + y^2 = 16. \qquad \text{Rectangular equation}$$

This means that the curve lies on a right circular cylinder of radius 4, centered about the z-axis. To locate the curve on this cylinder, you can use the third parametric equation $z = t$. In Figure 12.3, note that as t increases from 0 to 4π, the point (x, y, z) spirals up the cylinder to produce a **helix.** A real-life example of a helix is shown in the drawing at the lower left.

Try It | Exploration A | Open Exploration

In 1953 Francis Crick and James D. Watson discovered the double helix structure of DNA, which led to the $30 billion per year biotechnology industry.

Rotatable Graph

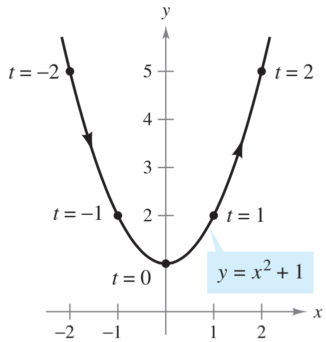
In Examples 1 and 2, you were given a vector-valued function and were asked to sketch the corresponding curve. The next two examples address the reverse problem—finding a vector-valued function to represent a given graph. Of course, if the graph is described parametrically, representation by a vector-valued function is straightforward. For instance, to represent the line in space given by

$$x = 2 + t, \quad y = 3t, \quad \text{and} \quad z = 4 - t$$

you can simply use the vector-valued function given by

$$\mathbf{r}(t) = (2 + t)\mathbf{i} + 3t\mathbf{j} + (4 - t)\mathbf{k}.$$

If a set of parametric equations for the graph is not given, the problem of representing the graph by a vector-valued function boils down to finding a set of parametric equations.



There are many ways to parametrize this graph. One way is to let $x = t$.
Figure 12.4

Editable Graph

EXAMPLE 3 Representing a Graph by a Vector-Valued Function

Represent the parabola given by $y = x^2 + 1$ by a vector-valued function.

Solution Although there are many ways to choose the parameter t, a natural choice is to let $x = t$. Then $y = t^2 + 1$ and you have

$$\mathbf{r}(t) = t\mathbf{i} + (t^2 + 1)\mathbf{j}.$$ Vector-valued function

Note in Figure 12.4 the orientation produced by this particular choice of parameter. Had you chosen $x = -t$ as the parameter, the curve would have been oriented in the opposite direction.

Try It Exploration A Exploration B

EXAMPLE 4 Representing a Graph by a Vector-Valued Function

Sketch the graph C represented by the intersection of the semiellipsoid

$$\frac{x^2}{12} + \frac{y^2}{24} + \frac{z^2}{4} = 1, \quad z \geq 0$$

and the parabolic cylinder $y = x^2$. Then, find a vector-valued function to represent the graph.

Solution The intersection of the two surfaces is shown in Figure 12.5. As in Example 3, a natural choice of parameter is $x = t$. For this choice, you can use the given equation $y = x^2$ to obtain $y = t^2$. Then, it follows that

$$\frac{z^2}{4} = 1 - \frac{x^2}{12} - \frac{y^2}{24} = 1 - \frac{t^2}{12} - \frac{t^4}{24} = \frac{24 - 2t^2 - t^4}{24}.$$

NOTE Curves in space can be specified in various ways. For instance, the curve in Example 4 is described as the intersection of two surfaces in space.

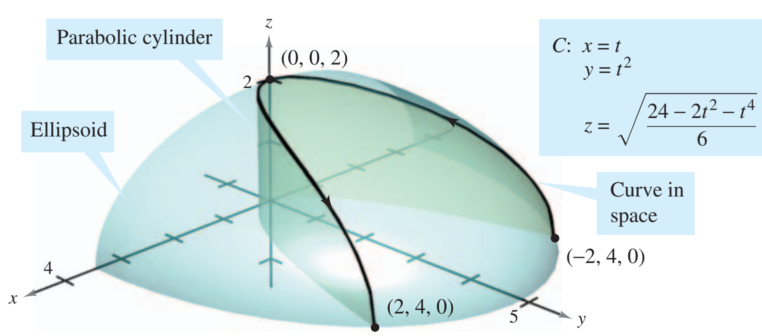
Because the curve lies above the xy-plane, you should choose the positive square root for z and obtain the following parametric equations.

$$x = t, \quad y = t^2, \quad \text{and} \quad z = \sqrt{\frac{24 - 2t^2 - t^4}{6}}$$

The resulting vector-valued function is

$$\mathbf{r}(t) = t\mathbf{i} + t^2\mathbf{j} + \sqrt{\frac{24 - 2t^2 - t^4}{6}}\,\mathbf{k}, \quad -2 \leq t \leq 2.$$ Vector-valued function

From the points $(-2, 4, 0)$ and $(2, 4, 0)$ shown in Figure 12.5, you can see that the curve is traced as t increases from -2 to 2.



The curve C is the intersection of the semiellipsoid and the parabolic cylinder.
Figure 12.5

Rotatable Graph

Try It Exploration A

Limits and Continuity

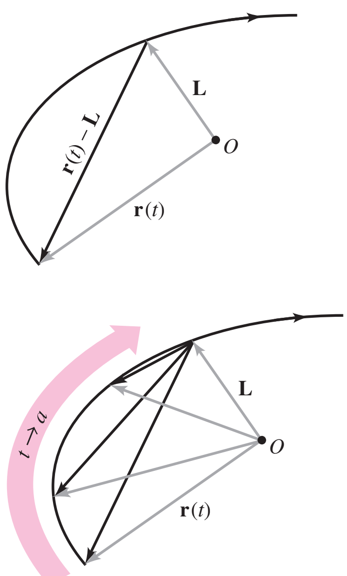
Many techniques and definitions used in the calculus of real-valued functions can be applied to vector-valued functions. For instance, you can add and subtract vector-valued functions, multiply a vector-valued function by a scalar, take the limit of a vector-valued function, differentiate a vector-valued function, and so on. The basic approach is to capitalize on the linearity of vector operations by extending the definitions on a component-by-component basis. For example, to add or subtract two vector-valued functions (in the plane), you can write

$$\begin{aligned} \mathbf{r}_1(t) + \mathbf{r}_2(t) &= [f_1(t)\mathbf{i} + g_1(t)\mathbf{j}] + [f_2(t)\mathbf{i} + g_2(t)\mathbf{j}] && \text{Sum} \\ &= [f_1(t) + f_2(t)]\mathbf{i} + [g_1(t) + g_2(t)]\mathbf{j} \\ \mathbf{r}_1(t) - \mathbf{r}_2(t) &= [f_1(t)\mathbf{i} + g_1(t)\mathbf{j}] - [f_2(t)\mathbf{i} + g_2(t)\mathbf{j}] && \text{Difference} \\ &= [f_1(t) - f_2(t)]\mathbf{i} + [g_1(t) - g_2(t)]\mathbf{j}. \end{aligned}$$

Similarly, to multiply and divide a vector-valued function by a scalar, you can write

$$\begin{aligned} c\mathbf{r}(t) &= c[f_1(t)\mathbf{i} + g_1(t)\mathbf{j}] && \text{Scalar multiplication} \\ &= cf_1(t)\mathbf{i} + cg_1(t)\mathbf{j} \\ \frac{\mathbf{r}(t)}{c} &= \frac{[f_1(t)\mathbf{i} + g_1(t)\mathbf{j}]}{c}, \quad c \neq 0 && \text{Scalar division} \\ &= \frac{f_1(t)}{c}\mathbf{i} + \frac{g_1(t)}{c}\mathbf{j}. \end{aligned}$$

This component-by-component extension of operations with real-valued functions to vector-valued functions is further illustrated in the following definition of the limit of a vector-valued function.



As t approaches a, $\mathbf{r}(t)$ approaches the limit $\mathbf{L}$. For the limit $\mathbf{L}$ to exist, it is not necessary that $\mathbf{r}(a)$ be defined or that $\mathbf{r}(a)$ be equal to $\mathbf{L}$.
Figure 12.6

Animation

Definition of the Limit of a Vector-Valued Function

1. If $\mathbf{r}$ is a vector-valued function such that $\mathbf{r}(t) = f(t)\mathbf{i} + g(t)\mathbf{j}$, then

$$\lim_{t\to a} \mathbf{r}(t) = \left[\lim_{t\to a} f(t)\right]\mathbf{i} + \left[\lim_{t\to a} g(t)\right]\mathbf{j} \qquad \text{Plane}$$

provided f and g have limits as $t \to a$.

2. If $\mathbf{r}$ is a vector-valued function such that $\mathbf{r}(t) = f(t)\mathbf{i} + g(t)\mathbf{j} + h(t)\mathbf{k}$, then

$$\lim_{t\to a} \mathbf{r}(t) = \left[\lim_{t\to a} f(t)\right]\mathbf{i} + \left[\lim_{t\to a} g(t)\right]\mathbf{j} + \left[\lim_{t\to a} h(t)\right]\mathbf{k} \qquad \text{Space}$$

provided f, g, and h have limits as $t \to a$.

If $\mathbf{r}(t)$ approaches the vector $\mathbf{L}$ as $t \to a$, the length of the vector $\mathbf{r}(t) - \mathbf{L}$ approaches 0. That is,

$$\|\mathbf{r}(t) - \mathbf{L}\| \to 0 \qquad \text{as} \qquad t \to a.$$

This is illustrated graphically in Figure 12.6. With this definition of the limit of a vector-valued function, you can develop vector versions of most of the limit theorems given in Chapter 1. For example, the limit of the sum of two vector-valued functions is the sum of their individual limits. Also, you can use the orientation of the curve $\mathbf{r}(t)$ to define one-sided limits of vector-valued functions. The next definition extends the notion of continuity to vector-valued functions.

Definition of Continuity of a Vector-Valued Function

A vector-valued function $\mathbf{r}$ is **continuous at the point** given by $t = a$ if the limit of $\mathbf{r}(t)$ exists as $t \to a$ and

$$\lim_{t\to a} \mathbf{r}(t) = \mathbf{r}(a).$$

A vector-valued function $\mathbf{r}$ is **continuous on an interval** I if it is continuous at every point in the interval.

From this definition, it follows that a vector-valued function is continuous at $t = a$ if and only if each of its component functions is continuous at $t = a$.

EXAMPLE 5 Continuity of Vector-Valued Functions

Discuss the continuity of the vector-valued function given by

$$\mathbf{r}(t) = t\mathbf{i} + a\mathbf{j} + (a^2 - t^2)\mathbf{k}$$ *a is a constant.*

at $t = 0$.

Solution As t approaches 0, the limit is

$$\begin{aligned}\lim_{t\to 0} \mathbf{r}(t) &= \left[\lim_{t\to 0} t\right]\mathbf{i} + \left[\lim_{t\to 0} a\right]\mathbf{j} + \left[\lim_{t\to 0} (a^2 - t^2)\right]\mathbf{k}\\ &= 0\mathbf{i} + a\mathbf{j} + a^2\mathbf{k}\\ &= a\mathbf{j} + a^2\mathbf{k}.\end{aligned}$$

Because

$$\begin{aligned}\mathbf{r}(0) &= (0)\mathbf{i} + (a)\mathbf{j} + (a^2)\mathbf{k}\\ &= a\mathbf{j} + a^2\mathbf{k}\end{aligned}$$

you can conclude that $\mathbf{r}$ is continuous at $t = 0$. By similar reasoning, you can conclude that the vector-valued function $\mathbf{r}$ is continuous at all real-number values of t.

Try It **Exploration A**

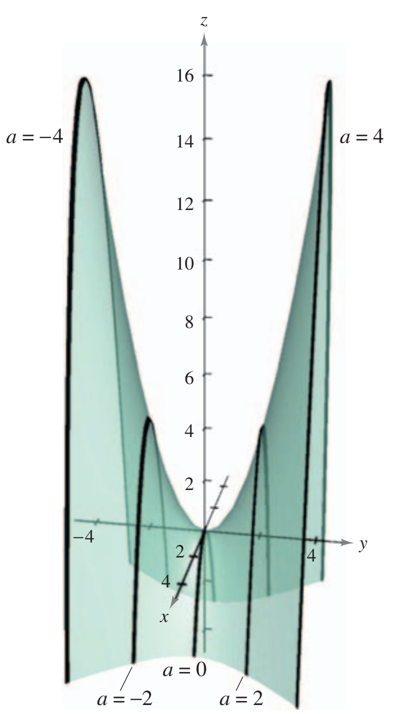
For each value of a, the curve represented by the vector-valued function in Example 5,

$$\mathbf{r}(t) = t\mathbf{i} + a\mathbf{j} + (a^2 - t^2)\mathbf{k}$$ *a is a constant.*

is a parabola. You can think of each parabola as the intersection of the vertical plane $y = a$ and the hyperbolic paraboloid

$$y^2 - x^2 = z$$

as shown in Figure 12.7.



For each value of a, the curve represented by the vector-valued function $\mathbf{r}(t) = t\mathbf{i} + a\mathbf{j} + (a^2 - t^2)\mathbf{k}$ is a parabola.
Figure 12.7

Rotatable Graph

TECHNOLOGY Almost any type of three-dimensional sketch is difficult to do by hand, but sketching curves in space is especially difficult. The problem is in trying to create the illusion of three dimensions. Graphing utilities use a variety of techniques to add "three-dimensionality" to graphs of space curves: one way is to show the curve on a surface, as in Figure 12.7.

Section 12.2 Differentiation and Integration of Vector-Valued Functions

- **Differentiate a vector-valued function.**
- **Integrate a vector-valued function.**

Differentiation of Vector-Valued Functions

In Sections 12.3–12.5, you will study several important applications involving the calculus of vector-valued functions. In preparation for that study, this section is devoted to the mechanics of differentiation and integration of vector-valued functions.

The definition of the derivative of a vector-valued function parallels that given for real-valued functions.

Definition of the Derivative of a Vector-Valued Function

The **derivative of a vector-valued function r** is defined by

$$\mathbf{r}'(t) = \lim_{\Delta t \to 0} \frac{\mathbf{r}(t + \Delta t) - \mathbf{r}(t)}{\Delta t}$$

for all t for which the limit exists. If $\mathbf{r}'(c)$ exists, then $\mathbf{r}$ is **differentiable at *c*.** If $\mathbf{r}'(c)$ exists for all c in an open interval I, then $\mathbf{r}$ is **differentiable on the interval *I*.** Differentiability of vector-valued functions can be extended to closed intervals by considering one-sided limits.

NOTE In addition to $\mathbf{r}'(t)$, other notations for the derivative of a vector-valued function are

$$D_t[\mathbf{r}(t)], \quad \frac{d}{dt}[\mathbf{r}(t)], \quad \text{and} \quad \frac{d\mathbf{r}}{dt}.$$

Differentiation of vector-valued functions can be done on a *component-by-component basis*. To see why this is true, consider the function given by

$$\mathbf{r}(t) = f(t)\mathbf{i} + g(t)\mathbf{j}.$$

Applying the definition of the derivative produces the following.

$$\begin{aligned}
\mathbf{r}'(t) &= \lim_{\Delta t \to 0} \frac{\mathbf{r}(t + \Delta t) - \mathbf{r}(t)}{\Delta t} \\
&= \lim_{\Delta t \to 0} \frac{f(t + \Delta t)\mathbf{i} + g(t + \Delta t)\mathbf{j} - f(t)\mathbf{i} - g(t)\mathbf{j}}{\Delta t} \\
&= \lim_{\Delta t \to 0} \left\{ \left[\frac{f(t + \Delta t) - f(t)}{\Delta t} \right] \mathbf{i} + \left[\frac{g(t + \Delta t) - g(t)}{\Delta t} \right] \mathbf{j} \right\} \\
&= \left\{ \lim_{\Delta t \to 0} \left[\frac{f(t + \Delta t) - f(t)}{\Delta t} \right] \right\} \mathbf{i} + \left\{ \lim_{\Delta t \to 0} \left[\frac{g(t + \Delta t) - g(t)}{\Delta t} \right] \right\} \mathbf{j} \\
&= f'(t)\mathbf{i} + g'(t)\mathbf{j}
\end{aligned}$$

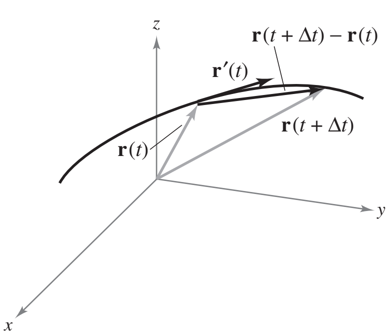


Figure 12.8

This important result is listed in the theorem on the next page. Note that the derivative of the vector-valued function $\mathbf{r}$ is itself a vector-valued function. You can see from Figure 12.8 that $\mathbf{r}'(t)$ is a vector tangent to the curve given by $\mathbf{r}(t)$ and pointing in the direction of increasing t-values.

THEOREM 12.1 Differentiation of Vector-Valued Functions

1. If $\mathbf{r}(t) = f(t)\mathbf{i} + g(t)\mathbf{j}$, where f and g are differentiable functions of t, then

$$\mathbf{r}'(t) = f'(t)\mathbf{i} + g'(t)\mathbf{j}. \qquad \text{Plane}$$

2. If $\mathbf{r}(t) = f(t)\mathbf{i} + g(t)\mathbf{j} + h(t)\mathbf{k}$, where f, g, and h are differentiable functions of t, then

$$\mathbf{r}'(t) = f'(t)\mathbf{i} + g'(t)\mathbf{j} + h'(t)\mathbf{k}. \qquad \text{Space}$$

EXAMPLE 1 Differentiation of Vector-Valued Functions

Find the derivative of each vector-valued function.

a. $\mathbf{r}(t) = t^2\mathbf{i} - 4\mathbf{j}$ **b.** $\mathbf{r}(t) = \dfrac{1}{t}\mathbf{i} + \ln t\mathbf{j} + e^{2t}\mathbf{k}$

Solution Differentiating on a component-by-component basis produces the following.

a. $\mathbf{r}'(t) = 2t\mathbf{i} - 0\mathbf{j}$

$= 2t\mathbf{i}$ Derivative

b. $\mathbf{r}'(t) = -\dfrac{1}{t^2}\mathbf{i} + \dfrac{1}{t}\mathbf{j} + 2e^{2t}\mathbf{k}$ Derivative

Try It Exploration A Open Exploration

Higher-order derivatives of vector-valued functions are obtained by successive differentiation of each component function.

EXAMPLE 2 Higher-Order Differentiation

For the vector-valued function given by $\mathbf{r}(t) = \cos t\mathbf{i} + \sin t\mathbf{j} + 2t\mathbf{k}$, find each of the following.

a. $\mathbf{r}'(t)$ **b.** $\mathbf{r}''(t)$

c. $\mathbf{r}'(t) \cdot \mathbf{r}''(t)$ **d.** $\mathbf{r}'(t) \times \mathbf{r}''(t)$

Solution

a. $\mathbf{r}'(t) = -\sin t\mathbf{i} + \cos t\mathbf{j} + 2\mathbf{k}$ First derivative

b. $\mathbf{r}''(t) = -\cos t\mathbf{i} - \sin t\mathbf{j} + 0\mathbf{k}$

$= -\cos t\mathbf{i} - \sin t\mathbf{j}$ Second derivative

c. $\mathbf{r}'(t) \cdot \mathbf{r}''(t) = \sin t \cos t - \sin t \cos t = 0$ Dot product

d. $\mathbf{r}'(t) \times \mathbf{r}''(t) = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ -\sin t & \cos t & 2 \\ -\cos t & -\sin t & 0 \end{vmatrix}$ Cross product

$$= \begin{vmatrix} \cos t & 2 \\ -\sin t & 0 \end{vmatrix}\mathbf{i} - \begin{vmatrix} -\sin t & 2 \\ -\cos t & 0 \end{vmatrix}\mathbf{j} + \begin{vmatrix} -\sin t & \cos t \\ -\cos t & -\sin t \end{vmatrix}\mathbf{k}$$

$$= 2\sin t\mathbf{i} - 2\cos t\mathbf{j} + \mathbf{k}$$

Note that the dot product in part (c) is a *real-valued* function, not a vector-valued function.

Try It Exploration A

The parametrization of the curve represented by the vector-valued function

$$\mathbf{r}(t) = f(t)\mathbf{i} + g(t)\mathbf{j} + h(t)\mathbf{k}$$

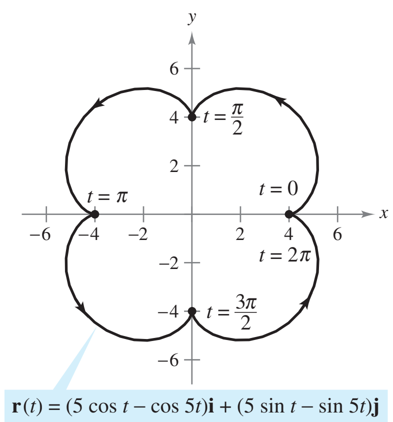
is **smooth on an open interval** I if f', g', and h' are continuous on I and $\mathbf{r}'(t) \neq \mathbf{0}$ for any value of t in the interval I.

EXAMPLE 3 Finding Intervals on Which a Curve Is Smooth

Find the intervals on which the epicycloid C given by

$$\mathbf{r}(t) = (5\cos t - \cos 5t)\mathbf{i} + (5\sin t - \sin 5t)\mathbf{j}, \quad 0 \le t \le 2\pi$$

is smooth.



The epicycloid is not smooth at the points where it intersects the axes.
Figure 12.9

Editable Graph

Solution The derivative of $\mathbf{r}$ is

$$\mathbf{r}'(t) = (-5\sin t + 5\sin 5t)\mathbf{i} + (5\cos t - 5\cos 5t)\mathbf{j}.$$

In the interval $[0, 2\pi]$, the only values of t for which

$$\mathbf{r}'(t) = 0\mathbf{i} + 0\mathbf{j}$$

are $t = 0, \pi/2, \pi, 3\pi/2$, and 2π. Therefore, you can conclude that C is smooth in the intervals

$$\left(0, \frac{\pi}{2}\right), \left(\frac{\pi}{2}, \pi\right), \left(\pi, \frac{3\pi}{2}\right), \quad \text{and} \quad \left(\frac{3\pi}{2}, 2\pi\right)$$

as shown in Figure 12.9.

Try It   Exploration A

NOTE In Figure 12.9, note that the curve is not smooth at points at which the curve makes abrupt changes in direction. Such points are called **cusps** or **nodes.**

Most of the differentiation rules in Chapter 2 have counterparts for vector-valued functions, and several are listed in the following theorem. Note that the theorem contains three versions of "product rules." Property 3 gives the derivative of the product of a real-valued function f and a vector-valued function $\mathbf{r}$, Property 4 gives the derivative of the dot product of two vector-valued functions, and Property 5 gives the derivative of the cross product of two vector-valued functions (in space). Note that Property 5 applies only to three-dimensional vector-valued functions, because the cross product is not defined for two-dimensional vectors.

THEOREM 12.2 Properties of the Derivative

Let $\mathbf{r}$ and $\mathbf{u}$ be differentiable vector-valued functions of t, let f be a differentiable real-valued function of t, and let c be a scalar.

1. $D_t[c\mathbf{r}(t)] = c\mathbf{r}'(t)$
2. $D_t[\mathbf{r}(t) \pm \mathbf{u}(t)] = \mathbf{r}'(t) \pm \mathbf{u}'(t)$
3. $D_t[f(t)\mathbf{r}(t)] = f(t)\mathbf{r}'(t) + f'(t)\mathbf{r}(t)$
4. $D_t[\mathbf{r}(t) \cdot \mathbf{u}(t)] = \mathbf{r}(t) \cdot \mathbf{u}'(t) + \mathbf{r}'(t) \cdot \mathbf{u}(t)$
5. $D_t[\mathbf{r}(t) \times \mathbf{u}(t)] = \mathbf{r}(t) \times \mathbf{u}'(t) + \mathbf{r}'(t) \times \mathbf{u}(t)$
6. $D_t[\mathbf{r}(f(t))] = \mathbf{r}'(f(t))f'(t)$
7. If $\mathbf{r}(t) \cdot \mathbf{r}(t) = c$, then $\mathbf{r}(t) \cdot \mathbf{r}'(t) = 0$.

Proof To prove Property 4, let

$$\mathbf{r}(t) = f_1(t)\mathbf{i} + g_1(t)\mathbf{j} \quad \text{and} \quad \mathbf{u}(t) = f_2(t)\mathbf{i} + g_2(t)\mathbf{j}$$

where f_1, f_2, g_1, and g_2 are differentiable functions of t. Then,

$$\mathbf{r}(t) \cdot \mathbf{u}(t) = f_1(t)f_2(t) + g_1(t)g_2(t)$$

and it follows that

$$\begin{aligned} D_t[\mathbf{r}(t) \cdot \mathbf{u}(t)] &= f_1(t)f_2'(t) + f_1'(t)f_2(t) + g_1(t)g_2'(t) + g_1'(t)g_2(t) \\ &= [f_1(t)f_2'(t) + g_1(t)g_2'(t)] + [f_1'(t)f_2(t) + g_1'(t)g_2(t)] \\ &= \mathbf{r}(t) \cdot \mathbf{u}'(t) + \mathbf{r}'(t) \cdot \mathbf{u}(t). \end{aligned}$$

Proofs of the other properties are left as exercises (see Exercises 73–77 and Exercise 80).

EXPLORATION

Let $\mathbf{r}(t) = \cos t\mathbf{i} + \sin t\mathbf{j}$. Sketch the graph of $\mathbf{r}(t)$. Explain why the graph is a circle of radius 1 centered at the origin. Calculate $\mathbf{r}(\pi/4)$ and $\mathbf{r}'(\pi/4)$. Position the vector $\mathbf{r}'(\pi/4)$ so that its initial point is at the terminal point of $\mathbf{r}(\pi/4)$. What do you observe? Show that $\mathbf{r}(t) \cdot \mathbf{r}(t)$ is constant and that $\mathbf{r}(t) \cdot \mathbf{r}'(t) = 0$ for all t. How does this example relate to Property 7 of Theorem 12.2?

EXAMPLE 4 Using Properties of the Derivative

For the vector-valued functions given by

$$\mathbf{r}(t) = \frac{1}{t}\mathbf{i} - \mathbf{j} + \ln t\mathbf{k} \quad \text{and} \quad \mathbf{u}(t) = t^2\mathbf{i} - 2t\mathbf{j} + \mathbf{k}$$

find

a. $D_t[\mathbf{r}(t) \cdot \mathbf{u}(t)]$ and **b.** $D_t[\mathbf{u}(t) \times \mathbf{u}'(t)]$.

Solution

a. Because $\mathbf{r}'(t) = -\dfrac{1}{t^2}\mathbf{i} + \dfrac{1}{t}\mathbf{k}$ and $\mathbf{u}'(t) = 2t\mathbf{i} - 2\mathbf{j}$, you have

$$\begin{aligned} D_t[\mathbf{r}(t) \cdot \mathbf{u}(t)] &= \mathbf{r}(t) \cdot \mathbf{u}'(t) + \mathbf{r}'(t) \cdot \mathbf{u}(t) \\ &= \left(\frac{1}{t}\mathbf{i} - \mathbf{j} + \ln t\mathbf{k}\right) \cdot (2t\mathbf{i} - 2\mathbf{j}) \\ &\quad + \left(-\frac{1}{t^2}\mathbf{i} + \frac{1}{t}\mathbf{k}\right) \cdot (t^2\mathbf{i} - 2t\mathbf{j} + \mathbf{k}) \\ &= 2 + 2 + (-1) + \frac{1}{t} \\ &= 3 + \frac{1}{t}. \end{aligned}$$

b. Because $\mathbf{u}'(t) = 2t\mathbf{i} - 2\mathbf{j}$ and $\mathbf{u}''(t) = 2\mathbf{i}$, you have

$$\begin{aligned} D_t[\mathbf{u}(t) \times \mathbf{u}'(t)] &= [\mathbf{u}(t) \times \mathbf{u}''(t)] + [\mathbf{u}'(t) \times \mathbf{u}'(t)] \\ &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ t^2 & -2t & 1 \\ 2 & 0 & 0 \end{vmatrix} + \mathbf{0} \\ &= \begin{vmatrix} -2t & 1 \\ 0 & 0 \end{vmatrix}\mathbf{i} - \begin{vmatrix} t^2 & 1 \\ 2 & 0 \end{vmatrix}\mathbf{j} + \begin{vmatrix} t^2 & -2t \\ 2 & 0 \end{vmatrix}\mathbf{k} \\ &= 0\mathbf{i} - (-2)\mathbf{j} + 4t\mathbf{k} \\ &= 2\mathbf{j} + 4t\mathbf{k}. \end{aligned}$$

Try It Exploration A Exploration B

NOTE Try reworking parts (a) and (b) in Example 4 by first forming the dot and cross products and then differentiating to see that you obtain the same results.

Integration of Vector-Valued Functions

The following definition is a rational consequence of the definition of the derivative of a vector-valued function.

Definition of Integration of Vector-Valued Functions

1. If $\mathbf{r}(t) = f(t)\mathbf{i} + g(t)\mathbf{j}$, where f and g are continuous on $[a, b]$, then the **indefinite integral (antiderivative)** of $\mathbf{r}$ is

$$\int \mathbf{r}(t)\,dt = \left[\int f(t)\,dt\right]\mathbf{i} + \left[\int g(t)\,dt\right]\mathbf{j} \qquad \text{Plane}$$

and its **definite integral** over the interval $a \le t \le b$ is

$$\int_a^b \mathbf{r}(t)\,dt = \left[\int_a^b f(t)\,dt\right]\mathbf{i} + \left[\int_a^b g(t)\,dt\right]\mathbf{j}.$$

2. If $\mathbf{r}(t) = f(t)\mathbf{i} + g(t)\mathbf{j} + h(t)\mathbf{k}$, where f, g, and h are continuous on $[a, b]$, then the **indefinite integral (antiderivative)** of $\mathbf{r}$ is

$$\int \mathbf{r}(t)\,dt = \left[\int f(t)\,dt\right]\mathbf{i} + \left[\int g(t)\,dt\right]\mathbf{j} + \left[\int h(t)\,dt\right]\mathbf{k} \qquad \text{Space}$$

and its **definite integral** over the interval $a \le t \le b$ is

$$\int_a^b \mathbf{r}(t)\,dt = \left[\int_a^b f(t)\,dt\right]\mathbf{i} + \left[\int_a^b g(t)\,dt\right]\mathbf{j} + \left[\int_a^b h(t)\,dt\right]\mathbf{k}.$$

The antiderivative of a vector-valued function is a family of vector-valued functions all differing by a constant vector $\mathbf{C}$. For instance, if $\mathbf{r}(t)$ is a three-dimensional vector-valued function, then for the indefinite integral $\int \mathbf{r}(t)\,dt$, you obtain three constants of integration

$$\int f(t)\,dt = F(t) + C_1, \qquad \int g(t)\,dt = G(t) + C_2, \qquad \int h(t)\,dt = H(t) + C_3$$

where $F'(t) = f(t)$, $G'(t) = g(t)$, and $H'(t) = h(t)$. These three *scalar* constants produce one *vector* constant of integration,

$$\begin{aligned}\int \mathbf{r}(t)\,dt &= [F(t) + C_1]\mathbf{i} + [G(t) + C_2]\mathbf{j} + [H(t) + C_3]\mathbf{k}\\ &= [F(t)\mathbf{i} + G(t)\mathbf{j} + H(t)\mathbf{k}] + [C_1\mathbf{i} + C_2\mathbf{j} + C_3\mathbf{k}]\\ &= \mathbf{R}(t) + \mathbf{C}\end{aligned}$$

where $\mathbf{R}'(t) = \mathbf{r}(t)$.

EXAMPLE 5 Integrating a Vector-Valued Function

Find the indefinite integral

$$\int (t\,\mathbf{i} + 3\mathbf{j})\,dt.$$

Solution Integrating on a component-by-component basis produces

$$\int (t\,\mathbf{i} + 3\mathbf{j})\,dt = \frac{t^2}{2}\mathbf{i} + 3t\mathbf{j} + \mathbf{C}.$$

Try It **Exploration A**

Example 6 shows how to evaluate the definite integral of a vector-valued function.

EXAMPLE 6 Definite Integral of a Vector-Valued Function

Evaluate the integral

$$\int_0^1 \mathbf{r}(t)\,dt = \int_0^1 \left(\sqrt[3]{t}\,\mathbf{i} + \frac{1}{t+1}\mathbf{j} + e^{-t}\mathbf{k}\right) dt.$$

Solution

$$\begin{aligned}\int_0^1 \mathbf{r}(t)\,dt &= \left(\int_0^1 t^{1/3}\,dt\right)\mathbf{i} + \left(\int_0^1 \frac{1}{t+1}\,dt\right)\mathbf{j} + \left(\int_0^1 e^{-t}\,dt\right)\mathbf{k}\\ &= \left[\left(\frac{3}{4}\right)t^{4/3}\right]_0^1 \mathbf{i} + \Big[\ln|t+1|\Big]_0^1 \mathbf{j} + \Big[-e^{-t}\Big]_0^1 \mathbf{k}\\ &= \frac{3}{4}\mathbf{i} + (\ln 2)\mathbf{j} + \left(1 - \frac{1}{e}\right)\mathbf{k}\end{aligned}$$

Try It Exploration A

As with real-valued functions, you can narrow the family of antiderivatives of a vector-valued function $\mathbf{r}'$ down to a single antiderivative by imposing an initial condition on the vector-valued function $\mathbf{r}$. This is demonstrated in the next example.

EXAMPLE 7 The Antiderivative of a Vector-Valued Function

Find the antiderivative of

$$\mathbf{r}'(t) = \cos 2t\mathbf{i} - 2\sin t\mathbf{j} + \frac{1}{1+t^2}\mathbf{k}$$

that satisfies the initial condition $\mathbf{r}(0) = 3\mathbf{i} - 2\mathbf{j} + \mathbf{k}$.

Solution

$$\begin{aligned}\mathbf{r}(t) &= \int \mathbf{r}'(t)\,dt\\ &= \left(\int \cos 2t\,dt\right)\mathbf{i} + \left(\int -2\sin t\,dt\right)\mathbf{j} + \left(\int \frac{1}{1+t^2}\,dt\right)\mathbf{k}\\ &= \left(\frac{1}{2}\sin 2t + C_1\right)\mathbf{i} + (2\cos t + C_2)\mathbf{j} + (\arctan t + C_3)\mathbf{k}\end{aligned}$$

Letting $t = 0$ and using the fact that $\mathbf{r}(0) = 3\mathbf{i} - 2\mathbf{j} + \mathbf{k}$, you have

$$\begin{aligned}\mathbf{r}(0) &= (0 + C_1)\mathbf{i} + (2 + C_2)\mathbf{j} + (0 + C_3)\mathbf{k}\\ &= 3\mathbf{i} + (-2)\mathbf{j} + \mathbf{k}.\end{aligned}$$

Equating corresponding components produces

$$C_1 = 3, \qquad 2 + C_2 = -2, \qquad \text{and} \qquad C_3 = 1.$$

So, the antiderivative that satisfies the given initial condition is

$$\mathbf{r}(t) = \left(\frac{1}{2}\sin 2t + 3\right)\mathbf{i} + (2\cos t - 4)\mathbf{j} + (\arctan t + 1)\mathbf{k}.$$

Try It Exploration A

Section 12.3 Velocity and Acceleration

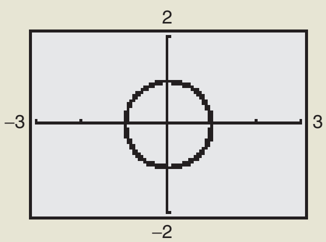
- **Describe the velocity and acceleration associated with a vector-valued function.**
- **Use a vector-valued function to analyze projectile motion.**

EXPLORATION

Exploring Velocity Consider the circle given by

$$\mathbf{r}(t) = (\cos \omega t)\mathbf{i} + (\sin \omega t)\mathbf{j}.$$

Use a graphing utility in *parametric* mode to graph this circle for several values of ω. How does ω affect the velocity of the terminal point as it traces out the curve? For a given value of ω, does the speed appear constant? Does the acceleration appear constant? Explain your reasoning.



Velocity and Acceleration

You are now ready to combine your study of parametric equations, curves, vectors, and vector-valued functions to form a model for motion along a curve. You will begin by looking at the motion of an object in the plane. (The motion of an object in space can be developed similarly.)

As an object moves along a curve in the plane, the coordinates x and y of its center of mass are each functions of time t. Rather than using the letters f and g to represent these two functions, it is convenient to write $x = x(t)$ and $y = y(t)$. So, the position vector $\mathbf{r}(t)$ takes the form

$$\mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j}. \qquad \text{Position vector}$$

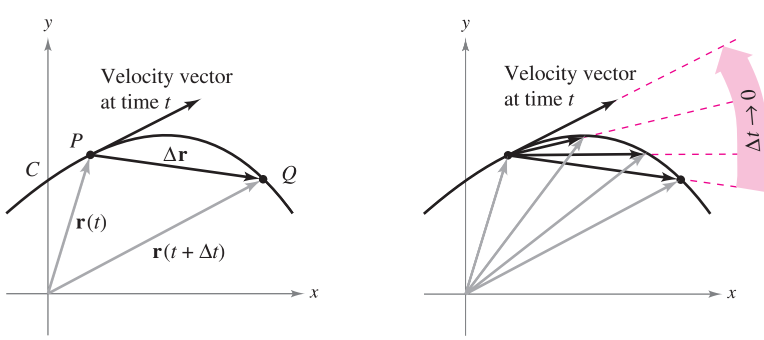
The beauty of this vector model for representing motion is that you can use the first and second derivatives of the vector-valued function $\mathbf{r}$ to find the object's velocity and acceleration. (Recall from the preceding chapter that velocity and acceleration are both vector quantities having magnitude and direction.) To find the velocity and acceleration vectors at a given time t, consider a point $Q(x(t + \Delta t), y(t + \Delta t))$ that is approaching the point $P(x(t), y(t))$ along the curve C given by $\mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j}$, as shown in Figure 12.10. As $\Delta t \to 0$, the direction of the vector $\overrightarrow{PQ}$ (denoted by $\Delta\mathbf{r}$) approaches the *direction of motion* at time t.

$$\Delta\mathbf{r} = \mathbf{r}(t + \Delta t) - \mathbf{r}(t)$$
$$\frac{\Delta\mathbf{r}}{\Delta t} = \frac{\mathbf{r}(t + \Delta t) - \mathbf{r}(t)}{\Delta t}$$
$$\lim_{\Delta t \to 0} \frac{\Delta\mathbf{r}}{\Delta t} = \lim_{\Delta t \to 0} \frac{\mathbf{r}(t + \Delta t) - \mathbf{r}(t)}{\Delta t}$$

If this limit exists, it is defined to be the **velocity vector** or **tangent vector** to the curve at point P. Note that this is the same limit used to define $\mathbf{r}'(t)$. So, the direction of $\mathbf{r}'(t)$ gives the direction of motion at time t. Moreover, the magnitude of the vector $\mathbf{r}'(t)$

$$\|\mathbf{r}'(t)\| = \|x'(t)\mathbf{i} + y'(t)\mathbf{j}\| = \sqrt{[x'(t)]^2 + [y'(t)]^2}$$

gives the **speed** of the object at time t. Similarly, you can use $\mathbf{r}''(t)$ to find acceleration, as indicated in the definitions at the top of page 849.



As $\Delta t \to 0$, $\dfrac{\Delta\mathbf{r}}{\Delta t}$ approaches the velocity vector.

Figure 12.10

Animation

Definitions of Velocity and Acceleration

If x and y are twice-differentiable functions of t, and $\mathbf{r}$ is a vector-valued function given by $\mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j}$, then the velocity vector, acceleration vector, and speed at time t are as follows.

$$\textbf{Velocity} = \mathbf{v}(t) = \mathbf{r}'(t) = x'(t)\mathbf{i} + y'(t)\mathbf{j}$$
$$\textbf{Acceleration} = \mathbf{a}(t) = \mathbf{r}''(t) = x''(t)\mathbf{i} + y''(t)\mathbf{j}$$
$$\textbf{Speed} = \|\mathbf{v}(t)\| = \|\mathbf{r}'(t)\| = \sqrt{[x'(t)]^2 + [y'(t)]^2}$$

For motion along a space curve, the definitions are similar. That is, if $\mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j} + z(t)\mathbf{k}$, you have

$$\textbf{Velocity} = \mathbf{v}(t) = \mathbf{r}'(t) = x'(t)\mathbf{i} + y'(t)\mathbf{j} + z'(t)\mathbf{k}$$
$$\textbf{Acceleration} = \mathbf{a}(t) = \mathbf{r}''(t) = x''(t)\mathbf{i} + y''(t)\mathbf{j} + z''(t)\mathbf{k}$$
$$\textbf{Speed} = \|\mathbf{v}(t)\| = \|\mathbf{r}'(t)\| = \sqrt{[x'(t)]^2 + [y'(t)]^2 + [z'(t)]^2}.$$

EXAMPLE 1 Finding Velocity and Acceleration Along a Plane Curve

NOTE In Example 1, note that the velocity and acceleration vectors are orthogonal at any point in time. This is characteristic of motion at a constant speed. (See Exercise 53.)

Find the velocity vector, speed, and acceleration vector of a particle that moves along the plane curve C described by

$$\mathbf{r}(t) = 2\sin\frac{t}{2}\mathbf{i} + 2\cos\frac{t}{2}\mathbf{j}. \qquad \text{Position vector}$$

Solution

The velocity vector is

$$\mathbf{v}(t) = \mathbf{r}'(t) = \cos\frac{t}{2}\mathbf{i} - \sin\frac{t}{2}\mathbf{j}. \qquad \text{Velocity vector}$$

The speed (at any time) is

$$\|\mathbf{r}'(t)\| = \sqrt{\cos^2\frac{t}{2} + \sin^2\frac{t}{2}} = 1. \qquad \text{Speed}$$

The acceleration vector is

$$\mathbf{a}(t) = \mathbf{r}''(t) = -\frac{1}{2}\sin\frac{t}{2}\mathbf{i} - \frac{1}{2}\cos\frac{t}{2}\mathbf{j}. \qquad \text{Acceleration vector}$$

Try It | Exploration A

The parametric equations for the curve in Example 1 are

$$x = 2\sin\frac{t}{2} \quad \text{and} \quad y = 2\cos\frac{t}{2}.$$

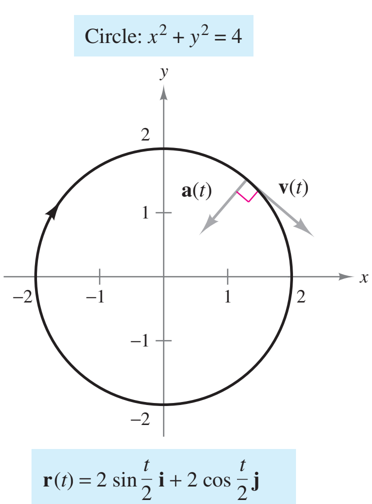
By eliminating the parameter t, you obtain the rectangular equation

$$x^2 + y^2 = 4. \qquad \text{Rectangular equation}$$

So, the curve is a circle of radius 2 centered at the origin, as shown in Figure 12.11. Because the velocity vector

$$\mathbf{v}(t) = \cos\frac{t}{2}\mathbf{i} - \sin\frac{t}{2}\mathbf{j}$$

has a constant magnitude but a changing direction as t increases, the particle moves around the circle at a constant speed.



The particle moves around the circle at a constant speed.
Figure 12.11

Editable Graph

EXAMPLE 2 Sketching Velocity and Acceleration Vectors in the Plane

Sketch the path of an object moving along the plane curve given by

$$\mathbf{r}(t) = (t^2 - 4)\mathbf{i} + t\mathbf{j} \qquad \text{Position vector}$$

and find the velocity and acceleration vectors when $t = 0$ and $t = 2$.

Solution Using the parametric equations $x = t^2 - 4$ and $y = t$, you can determine that the curve is a parabola given by $x = y^2 - 4$, as shown in Figure 12.12. The velocity vector (at any time) is

$$\mathbf{v}(t) = \mathbf{r}'(t) = 2t\mathbf{i} + \mathbf{j} \qquad \text{Velocity vector}$$

and the acceleration vector (at any time) is

$$\mathbf{a}(t) = \mathbf{r}''(t) = 2\mathbf{i}. \qquad \text{Acceleration vector}$$

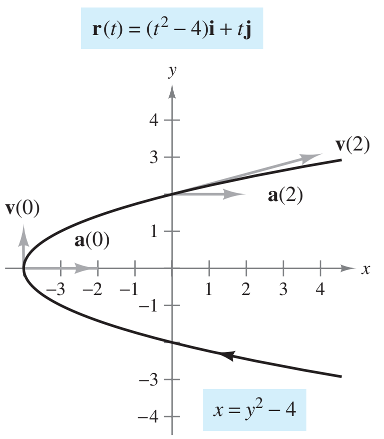
When $t = 0$, the velocity and acceleration vectors are given by

$$\mathbf{v}(0) = 2(0)\mathbf{i} + \mathbf{j} = \mathbf{j} \quad \text{and} \quad \mathbf{a}(0) = 2\mathbf{i}.$$

When $t = 2$, the velocity and acceleration vectors are given by

$$\mathbf{v}(2) = 2(2)\mathbf{i} + \mathbf{j} = 4\mathbf{i} + \mathbf{j} \quad \text{and} \quad \mathbf{a}(2) = 2\mathbf{i}.$$

Try It Exploration A

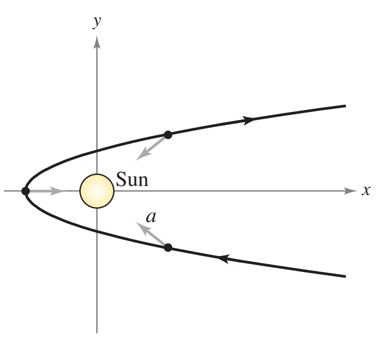


At each point on the curve, the acceleration vector points to the right.
Figure 12.12

Editable Graph

For the object moving along the path shown in Figure 12.12, note that the acceleration vector is constant (it has a magnitude of 2 and points to the right). This implies that the speed of the object is decreasing as the object moves toward the vertex of the parabola, and the speed is increasing as the object moves away from the vertex of the parabola.

This type of motion is *not* characteristic of comets that travel on parabolic paths through our solar system. For such comets, the acceleration vector always points to the origin (the sun), which implies that the comet's speed increases as it approaches the vertex of the path and decreases as it moves away from the vertex. (See Figure 12.13.)



At each point in the comet's orbit, the acceleration vector points toward the sun.
Figure 12.13

Animation

EXAMPLE 3 Sketching Velocity and Acceleration Vectors in Space

Sketch the path of an object moving along the space curve C given by

$$\mathbf{r}(t) = t\mathbf{i} + t^3\mathbf{j} + 3t\mathbf{k}, \quad t \geq 0 \qquad \text{Position vector}$$

and find the velocity and acceleration vectors when $t = 1$.

Solution Using the parametric equations $x = t$ and $y = t^3$, you can determine that the path of the object lies on the cubic cylinder given by $y = x^3$. Moreover, because $z = 3t$, the object starts at $(0, 0, 0)$ and moves upward as t increases, as shown in Figure 12.14. Because $\mathbf{r}(t) = t\mathbf{i} + t^3\mathbf{j} + 3t\mathbf{k}$, you have

$$\mathbf{v}(t) = \mathbf{r}'(t) = \mathbf{i} + 3t^2\mathbf{j} + 3\mathbf{k} \qquad \text{Velocity vector}$$

and

$$\mathbf{a}(t) = \mathbf{r}''(t) = 6t\mathbf{j}. \qquad \text{Acceleration vector}$$

When $t = 1$, the velocity and acceleration vectors are given by

$$\mathbf{v}(1) = \mathbf{r}'(1) = \mathbf{i} + 3\mathbf{j} + 3\mathbf{k} \quad \text{and} \quad \mathbf{a}(1) = \mathbf{r}''(1) = 6\mathbf{j}.$$

Try It Exploration A Open Exploration

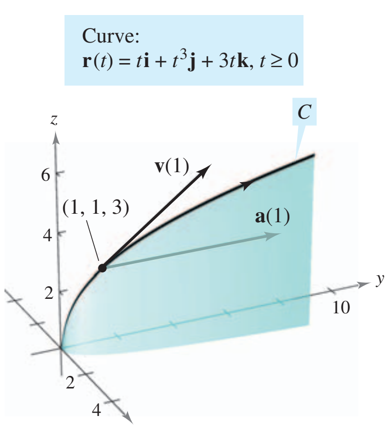


Figure 12.14

Rotatable Graph

So far in this section, you have concentrated on finding the velocity and acceleration by differentiating the position function. Many practical applications involve the reverse problem—finding the position function for a given velocity or acceleration. This is demonstrated in the next example.

EXAMPLE 4 Finding a Position Function by Integration

An object starts from rest at the point $P(1, 2, 0)$ and moves with an acceleration of

$$\mathbf{a}(t) = \mathbf{j} + 2\mathbf{k} \qquad \text{Acceleration vector}$$

where $\|\mathbf{a}(t)\|$ is measured in feet per second per second. Find the location of the object after $t = 2$ seconds.

Solution From the description of the object's motion, you can deduce the following *initial conditions*. Because the object starts from rest, you have

$$\mathbf{v}(0) = \mathbf{0}.$$

Moreover, because the object starts at the point $(x, y, z) = (1, 2, 0)$, you have

$$\begin{aligned}\mathbf{r}(0) &= x(0)\mathbf{i} + y(0)\mathbf{j} + z(0)\mathbf{k}\\ &= 1\mathbf{i} + 2\mathbf{j} + 0\mathbf{k}\\ &= \mathbf{i} + 2\mathbf{j}.\end{aligned}$$

To find the position function, you should integrate twice, each time using one of the initial conditions to solve for the constant of integration. The velocity vector is

$$\begin{aligned}\mathbf{v}(t) = \int \mathbf{a}(t)\,dt &= \int (\mathbf{j} + 2\mathbf{k})\,dt\\ &= t\mathbf{j} + 2t\mathbf{k} + \mathbf{C}\end{aligned}$$

where $\mathbf{C} = C_1\mathbf{i} + C_2\mathbf{j} + C_3\mathbf{k}$. Letting $t = 0$ and applying the initial condition $\mathbf{v}(0) = \mathbf{0}$, you obtain

$$\mathbf{v}(0) = C_1\mathbf{i} + C_2\mathbf{j} + C_3\mathbf{k} = \mathbf{0} \quad \Rightarrow \quad C_1 = C_2 = C_3 = 0.$$

So, the *velocity* at any time t is

$$\mathbf{v}(t) = t\mathbf{j} + 2t\mathbf{k}. \qquad \text{Velocity vector}$$

Integrating once more produces

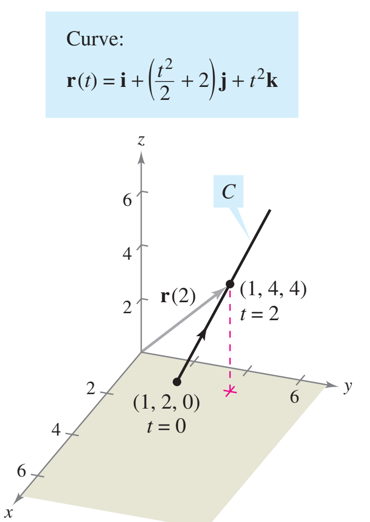
$$\begin{aligned}\mathbf{r}(t) = \int \mathbf{v}(t)\,dt &= \int (t\mathbf{j} + 2t\mathbf{k})\,dt\\ &= \frac{t^2}{2}\mathbf{j} + t^2\mathbf{k} + \mathbf{C}\end{aligned}$$

where $\mathbf{C} = C_4\mathbf{i} + C_5\mathbf{j} + C_6\mathbf{k}$. Letting $t = 0$ and applying the initial condition $\mathbf{r}(0) = \mathbf{i} + 2\mathbf{j}$, you have

$$\mathbf{r}(0) = C_4\mathbf{i} + C_5\mathbf{j} + C_6\mathbf{k} = \mathbf{i} + 2\mathbf{j} \quad \Rightarrow \quad C_4 = 1, C_5 = 2, C_6 = 0.$$

So, the *position* vector is

$$\mathbf{r}(t) = \mathbf{i} + \left(\frac{t^2}{2} + 2\right)\mathbf{j} + t^2\mathbf{k}. \qquad \text{Position vector}$$



The object takes 2 seconds to move from point $(1, 2, 0)$ to point $(1, 4, 4)$ along C.
Figure 12.15

The location of the object after $t = 2$ seconds is given by $\mathbf{r}(2) = \mathbf{i} + 4\mathbf{j} + 4\mathbf{k}$, as shown in Figure 12.15.

Try It **Exploration A**

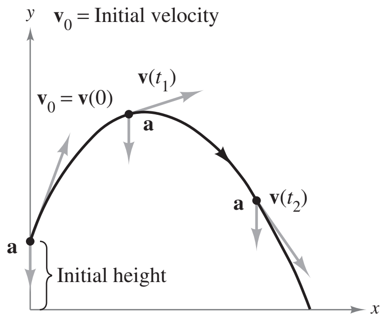


Figure 12.16

Projectile Motion

You now have the machinery to derive the parametric equations for the path of a projectile. Assume that gravity is the only force acting on the projectile after it is launched. So, the motion occurs in a vertical plane, which can be represented by the xy-coordinate system with the origin as a point on Earth's surface, as shown in Figure 12.16. For a projectile of mass m, the force due to gravity is

$$\mathbf{F} = -mg\mathbf{j} \qquad \text{Force due to gravity}$$

where the gravitational constant is $g = 32$ feet per second per second, or 9.81 meters per second per second. By **Newton's Second Law of Motion,** this same force produces an acceleration $\mathbf{a} = \mathbf{a}(t)$, and satisfies the equation $\mathbf{F} = m\mathbf{a}$. Consequently, the acceleration of the projectile is given by $m\mathbf{a} = -mg\mathbf{j}$, which implies that

$$\mathbf{a} = -g\mathbf{j}. \qquad \text{Acceleration of projectile}$$

EXAMPLE 5 Derivation of the Position Function for a Projectile

A projectile of mass m is launched from an initial position $\mathbf{r}_0$ with an initial velocity $\mathbf{v}_0$. Find its position vector as a function of time.

Solution Begin with the acceleration $\mathbf{a}(t) = -g\mathbf{j}$ and integrate twice.

$$\mathbf{v}(t) = \int \mathbf{a}(t)\,dt = \int -g\mathbf{j}\,dt = -gt\mathbf{j} + \mathbf{C}_1$$

$$\mathbf{r}(t) = \int \mathbf{v}(t)\,dt = \int (-gt\mathbf{j} + \mathbf{C}_1)\,dt = -\frac{1}{2}gt^2\mathbf{j} + \mathbf{C}_1 t + \mathbf{C}_2$$

You can use the facts that $\mathbf{v}(0) = \mathbf{v}_0$ and $\mathbf{r}(0) = \mathbf{r}_0$ to solve for the constant vectors $\mathbf{C}_1$ and $\mathbf{C}_2$. Doing this produces $\mathbf{C}_1 = \mathbf{v}_0$ and $\mathbf{C}_2 = \mathbf{r}_0$. Therefore, the position vector is

$$\mathbf{r}(t) = -\frac{1}{2}gt^2\mathbf{j} + t\mathbf{v}_0 + \mathbf{r}_0. \qquad \text{Position vector}$$

Exploration A

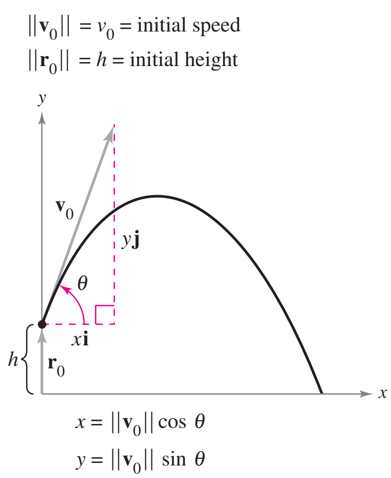


Figure 12.17

In many projectile problems, the constant vectors $\mathbf{r}_0$ and $\mathbf{v}_0$ are not given explicitly. Often you are given the initial height h, the initial speed v_0, and the angle θ at which the projectile is launched, as shown in Figure 12.17. From the given height, you can deduce that $\mathbf{r}_0 = h\mathbf{j}$. Because the speed gives the magnitude of the initial velocity, it follows that $v_0 = \|\mathbf{v}_0\|$ and you can write

$$\begin{aligned}\mathbf{v}_0 &= x\mathbf{i} + y\mathbf{j}\\ &= (\|\mathbf{v}_0\|\cos\theta)\mathbf{i} + (\|\mathbf{v}_0\|\sin\theta)\mathbf{j}\\ &= v_0\cos\theta\mathbf{i} + v_0\sin\theta\mathbf{j}.\end{aligned}$$

So, the position vector can be written in the form

$$\boxed{\mathbf{r}(t) = -\frac{1}{2}gt^2\mathbf{j} + t\mathbf{v}_0 + \mathbf{r}_0} \qquad \text{Position vector}$$

$$= -\frac{1}{2}gt^2\mathbf{j} + tv_0\cos\theta\mathbf{i} + tv_0\sin\theta\mathbf{j} + h\mathbf{j}$$

$$= (v_0\cos\theta)t\mathbf{i} + \left[h + (v_0\sin\theta)t - \frac{1}{2}gt^2\right]\mathbf{j}.$$

THEOREM 12.3 Position Function for a Projectile

Neglecting air resistance, the path of a projectile launched from an initial height h with initial speed v_0 and angle of elevation θ is described by the vector function

$$\mathbf{r}(t) = (v_0 \cos \theta)t\mathbf{i} + \left[h + (v_0 \sin \theta)t - \frac{1}{2}gt^2\right]\mathbf{j}$$

where g is the gravitational constant.

EXAMPLE 6 Describing the Path of a Baseball

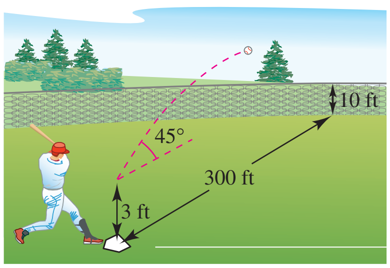


Figure 12.18

A baseball is hit 3 feet above ground level at 100 feet per second and at an angle of 45° with respect to the ground, as shown in Figure 12.18. Find the maximum height reached by the baseball. Will it clear a 10-foot-high fence located 300 feet from home plate?

Solution You are given $h = 3$, $v_0 = 100$, and $\theta = 45°$. So, using $g = 32$ feet per second per second produces

$$\begin{aligned}\mathbf{r}(t) &= \left(100 \cos \frac{\pi}{4}\right)t\mathbf{i} + \left[3 + \left(100 \sin \frac{\pi}{4}\right)t - 16t^2\right]\mathbf{j}\\ &= \left(50\sqrt{2}\,t\right)\mathbf{i} + \left(3 + 50\sqrt{2}\,t - 16t^2\right)\mathbf{j}\\ \mathbf{v}(t) &= \mathbf{r}'(t) = 50\sqrt{2}\,\mathbf{i} + \left(50\sqrt{2} - 32t\right)\mathbf{j}.\end{aligned}$$

The maximum height occurs when

$$y'(t) = 50\sqrt{2} - 32t = 0$$

which implies that

$$\begin{aligned}t &= \frac{25\sqrt{2}}{16}\\ &\approx 2.21 \text{ seconds.}\end{aligned}$$

So, the maximum height reached by the ball is

$$\begin{aligned}y &= 3 + 50\sqrt{2}\left(\frac{25\sqrt{2}}{16}\right) - 16\left(\frac{25\sqrt{2}}{16}\right)^2\\ &= \frac{649}{8}\\ &\approx 81 \text{ feet.}\end{aligned}$$

Maximum height when $t \approx 2.21$ seconds

The ball is 300 feet from where it was hit when

$$300 = x(t) = 50\sqrt{2}\,t.$$

Solving this equation for t produces $t = 3\sqrt{2} \approx 4.24$ seconds. At this time, the height of the ball is

$$\begin{aligned}y &= 3 + 50\sqrt{2}\left(3\sqrt{2}\right) - 16\left(3\sqrt{2}\right)^2\\ &= 303 - 288\\ &= 15 \text{ feet.}\end{aligned}$$

Height when $t \approx 4.24$ seconds

Therefore, the ball clears the 10-foot fence for a home run.

Try It

Exploration A

Section 12.4 Tangent Vectors and Normal Vectors

- **Find a unit tangent vector at a point on a space curve.**
- **Find the tangential and normal components of acceleration.**

Tangent Vectors and Normal Vectors

In the preceding section, you learned that the velocity vector points in the direction of motion. This observation leads to the following definition, which applies to any smooth curve—not just to those for which the parameter represents time.

Definition of Unit Tangent Vector

Let C be a smooth curve represented by $\mathbf{r}$ on an open interval I. The **unit tangent vector** $\mathbf{T}(t)$ at t is defined to be

$$\mathbf{T}(t) = \frac{\mathbf{r}'(t)}{\|\mathbf{r}'(t)\|}, \quad \mathbf{r}'(t) \neq \mathbf{0}.$$

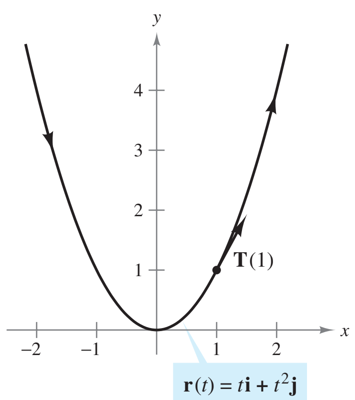
Recall that a curve is *smooth* on an interval if $\mathbf{r}'$ is continuous and nonzero on the interval. So, "smoothness" is sufficient to guarantee that a curve has a unit tangent vector.

EXAMPLE 1 Finding the Unit Tangent Vector

Find the unit tangent vector to the curve given by

$$\mathbf{r}(t) = t\mathbf{i} + t^2\mathbf{j}$$

when $t = 1$.



The direction of the unit tangent vector depends on the orientation of the curve.
Figure 12.19

Editable Graph

Solution The derivative of $\mathbf{r}(t)$ is

$$\mathbf{r}'(t) = \mathbf{i} + 2t\mathbf{j}. \qquad \text{Derivative of } \mathbf{r}(t)$$

So, the unit tangent vector is

$$\mathbf{T}(t) = \frac{\mathbf{r}'(t)}{\|\mathbf{r}'(t)\|} \qquad \text{Definition of } \mathbf{T}(t)$$

$$= \frac{1}{\sqrt{1 + 4t^2}}(\mathbf{i} + 2t\mathbf{j}). \qquad \text{Substitute for } \mathbf{r}'(t).$$

When $t = 1$, the unit tangent vector is

$$\mathbf{T}(1) = \frac{1}{\sqrt{5}}(\mathbf{i} + 2\mathbf{j})$$

as shown in Figure 12.19.

Try It Exploration A

NOTE In Example 1, note that the direction of the unit tangent vector depends on the orientation of the curve. For instance, if the parabola in Figure 12.19 were given by

$$\mathbf{r}(t) = -(t - 2)\mathbf{i} + (t - 2)^2\mathbf{j},$$

$\mathbf{T}(1)$ would still represent the unit tangent vector at the point $(1, 1)$, but it would point in the opposite direction. Try verifying this.

The **tangent line to a curve** at a point is the line passing through the point and parallel to the unit tangent vector. In Example 2, the unit tangent vector is used to find the tangent line at a point on a helix.

EXAMPLE 2 Finding the Tangent Line at a Point on a Curve

Find $\mathbf{T}(t)$ and then find a set of parametric equations for the tangent line to the helix given by

$$\mathbf{r}(t) = 2\cos t\mathbf{i} + 2\sin t\mathbf{j} + t\mathbf{k}$$

at the point corresponding to $t = \pi/4$.

Solution The derivative of $\mathbf{r}(t)$ is $\mathbf{r}'(t) = -2\sin t\mathbf{i} + 2\cos t\mathbf{j} + \mathbf{k}$, which implies that $\|\mathbf{r}'(t)\| = \sqrt{4\sin^2 t + 4\cos^2 t + 1} = \sqrt{5}$. Therefore, the unit tangent vector is

$$\mathbf{T}(t) = \frac{\mathbf{r}'(t)}{\|\mathbf{r}'(t)\|}$$

$$= \frac{1}{\sqrt{5}}(-2\sin t\mathbf{i} + 2\cos t\mathbf{j} + \mathbf{k}). \qquad \text{Unit tangent vector}$$

When $t = \pi/4$, the unit tangent vector is

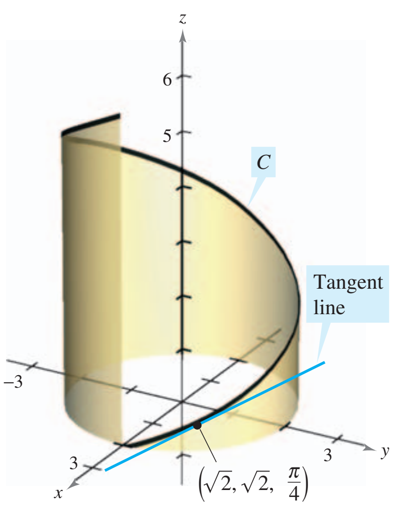
$$\mathbf{T}\left(\frac{\pi}{4}\right) = \frac{1}{\sqrt{5}}\left(-2\frac{\sqrt{2}}{2}\mathbf{i} + 2\frac{\sqrt{2}}{2}\mathbf{j} + \mathbf{k}\right)$$

$$= \frac{1}{\sqrt{5}}\left(-\sqrt{2}\mathbf{i} + \sqrt{2}\mathbf{j} + \mathbf{k}\right).$$

Using the direction numbers $a = -\sqrt{2}$, $b = \sqrt{2}$, and $c = 1$, and the point $(x_1, y_1, z_1) = (\sqrt{2}, \sqrt{2}, \pi/4)$, you can obtain the following parametric equations (given with parameter s).

$$x = x_1 + as = \sqrt{2} - \sqrt{2}s$$
$$y = y_1 + bs = \sqrt{2} + \sqrt{2}s$$
$$z = z_1 + cs = \frac{\pi}{4} + s$$

This tangent line is shown in Figure 12.20.



The tangent line to a curve at a point is determined by the unit tangent vector at the point.
Figure 12.20

Rotatable Graph

Try It | Exploration A

In Example 2, there are infinitely many vectors that are orthogonal to the tangent vector $\mathbf{T}(t)$. One of these is the vector $\mathbf{T}'(t)$. This follows from Property 7 of Theorem 12.2. That is,

$$\mathbf{T}(t)\cdot\mathbf{T}(t) = \|\mathbf{T}(t)\|^2 = 1 \quad \Longrightarrow \quad \mathbf{T}(t)\cdot\mathbf{T}'(t) = 0.$$

By normalizing the vector $\mathbf{T}'(t)$, you obtain a special vector called the **principal unit normal vector,** as indicated in the following definition.

Definition of Principal Unit Normal Vector

Let C be a smooth curve represented by $\mathbf{r}$ on an open interval I. If $\mathbf{T}'(t) \neq \mathbf{0}$, then the **principal unit normal vector** at t is defined to be

$$\mathbf{N}(t) = \frac{\mathbf{T}'(t)}{\|\mathbf{T}'(t)\|}.$$

EXAMPLE 3 Finding the Principal Unit Normal Vector

Find $\mathbf{N}(t)$ and $\mathbf{N}(1)$ for the curve represented by

$$\mathbf{r}(t) = 3t\mathbf{i} + 2t^2\mathbf{j}.$$

Solution By differentiating, you obtain

$$\mathbf{r}'(t) = 3\mathbf{i} + 4t\mathbf{j} \qquad \text{and} \qquad \|\mathbf{r}'(t)\| = \sqrt{9 + 16t^2}$$

which implies that the unit tangent vector is

$$\begin{aligned}\mathbf{T}(t) &= \frac{\mathbf{r}'(t)}{\|\mathbf{r}'(t)\|}\\ &= \frac{1}{\sqrt{9 + 16t^2}}(3\mathbf{i} + 4t\mathbf{j}).\end{aligned}$$

Unit tangent vector

Using Theorem 12.2, differentiate $\mathbf{T}(t)$ with respect to t to obtain

$$\begin{aligned}\mathbf{T}'(t) &= \frac{1}{\sqrt{9 + 16t^2}}(4\mathbf{j}) - \frac{16t}{(9 + 16t^2)^{3/2}}(3\mathbf{i} + 4t\mathbf{j})\\ &= \frac{12}{(9 + 16t^2)^{3/2}}(-4t\mathbf{i} + 3\mathbf{j})\\ \|\mathbf{T}'(t)\| &= 12\sqrt{\frac{9 + 16t^2}{(9 + 16t^2)^3}} = \frac{12}{9 + 16t^2}.\end{aligned}$$

Therefore, the principal unit normal vector is

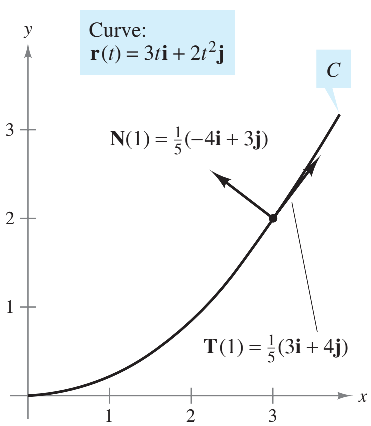
$$\begin{aligned}\mathbf{N}(t) &= \frac{\mathbf{T}'(t)}{\|\mathbf{T}'(t)\|}\\ &= \frac{1}{\sqrt{9 + 16t^2}}(-4t\mathbf{i} + 3\mathbf{j}).\end{aligned}$$

Principal unit normal vector

When $t = 1$, the principal unit normal vector is

$$\mathbf{N}(1) = \frac{1}{5}(-4\mathbf{i} + 3\mathbf{j})$$

as shown in Figure 12.21.



The principal unit normal vector points toward the concave side of the curve.
Figure 12.21

Editable Graph

Try It

Exploration A

The principal unit normal vector can be difficult to evaluate algebraically. For plane curves, you can simplify the algebra by finding

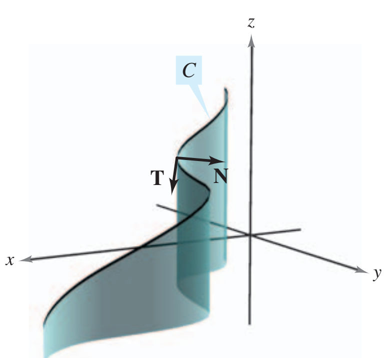
$$\mathbf{T}(t) = x(t)\mathbf{i} + y(t)\mathbf{j}$$

Unit tangent vector

and observing that $\mathbf{N}(t)$ must be either

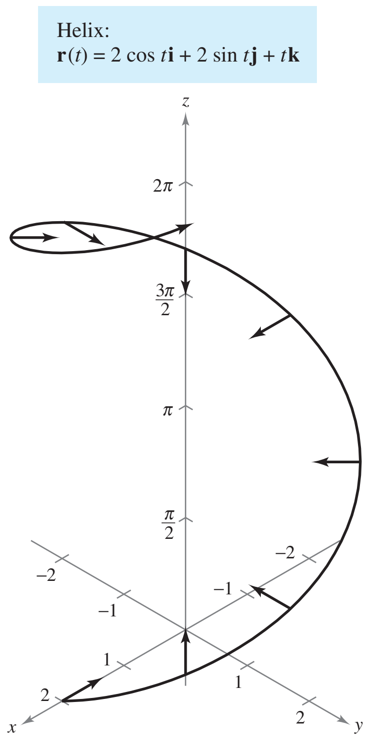
$$\mathbf{N}_1(t) = y(t)\mathbf{i} - x(t)\mathbf{j} \qquad \text{or} \qquad \mathbf{N}_2(t) = -y(t)\mathbf{i} + x(t)\mathbf{j}.$$

Because $\sqrt{[x(t)]^2 + [y(t)]^2} = 1$, it follows that both $\mathbf{N}_1(t)$ and $\mathbf{N}_2(t)$ are unit normal vectors. The *principal* unit normal vector $\mathbf{N}$ is the one that points toward the concave side of the curve, as shown in Figure 12.21 (see Exercise 86). This also holds for curves in space. That is, for an object moving along a curve C in space, the vector $\mathbf{T}(t)$ points in the direction the object is moving, whereas the vector $\mathbf{N}(t)$ is orthogonal to $\mathbf{T}(t)$ and points in the direction in which the object is turning, as shown in Figure 12.22.



At any point on a curve, a unit normal vector is orthogonal to the unit tangent vector. The *principal* unit normal vector points in the direction in which the curve is turning.
Figure 12.22

Rotatable Graph



$\mathbf{N}(t)$ is horizontal and points toward the z-axis.
Figure 12.23

EXAMPLE 4 Finding the Principal Unit Normal Vector

Find the principal unit normal vector for the helix given by

$$\mathbf{r}(t) = 2\cos t\mathbf{i} + 2\sin t\mathbf{j} + t\mathbf{k}.$$

Solution From Example 2, you know that the unit tangent vector is

$$\mathbf{T}(t) = \frac{1}{\sqrt{5}}(-2\sin t\mathbf{i} + 2\cos t\mathbf{j} + \mathbf{k}). \qquad \text{Unit tangent vector}$$

So, $\mathbf{T}'(t)$ is given by

$$\mathbf{T}'(t) = \frac{1}{\sqrt{5}}(-2\cos t\mathbf{i} - 2\sin t\mathbf{j}).$$

Because $\|\mathbf{T}'(t)\| = 2/\sqrt{5}$, it follows that the principal unit normal vector is

$$\begin{aligned}\mathbf{N}(t) &= \frac{\mathbf{T}'(t)}{\|\mathbf{T}'(t)\|}\\ &= \frac{1}{2}(-2\cos t\mathbf{i} - 2\sin t\mathbf{j})\\ &= -\cos t\mathbf{i} - \sin t\mathbf{j}. \qquad \text{Principal unit normal vector}\end{aligned}$$

Note that this vector is horizontal and points toward the z-axis, as shown in Figure 12.23.

Try It | Exploration A

Tangential and Normal Components of Acceleration

Let's return to the problem of describing the motion of an object along a curve. In the preceding section, you saw that for an object traveling at a *constant speed*, the velocity and acceleration vectors are perpendicular. This seems reasonable, because the speed would not be constant if any acceleration were acting in the direction of motion. You can verify this observation by noting that

$$\mathbf{r}''(t) \cdot \mathbf{r}'(t) = 0$$

if $\|\mathbf{r}'(t)\|$ is a constant. (See Property 7 of Theorem 12.2.)

However, for an object traveling at a *variable speed*, the velocity and acceleration vectors are not necessarily perpendicular. For instance, you saw that the acceleration vector for a projectile always points down, regardless of the direction of motion.

In general, part of the acceleration (the tangential component) acts in the line of motion, and part (the normal component) acts perpendicular to the line of motion. In order to determine these two components, you can use the unit vectors $\mathbf{T}(t)$ and $\mathbf{N}(t)$, which serve in much the same way as do $\mathbf{i}$ and $\mathbf{j}$ in representing vectors in the plane. The following theorem states that the acceleration vector lies in the plane determined by $\mathbf{T}(t)$ and $\mathbf{N}(t)$.

THEOREM 12.4 Acceleration Vector

If $\mathbf{r}(t)$ is the position vector for a smooth curve C and $\mathbf{N}(t)$ exists, then the acceleration vector $\mathbf{a}(t)$ lies in the plane determined by $\mathbf{T}(t)$ and $\mathbf{N}(t)$.

Proof To simplify the notation, write $\mathbf{T}$ for $\mathbf{T}(t)$, $\mathbf{T}'$ for $\mathbf{T}'(t)$, and so on. Because $\mathbf{T} = \mathbf{r}'/\|\mathbf{r}'\| = \mathbf{v}/\|\mathbf{v}\|$, it follows that

$$\mathbf{v} = \|\mathbf{v}\|\mathbf{T}.$$

By differentiating, you obtain

$$\begin{aligned}
\mathbf{a} = \mathbf{v}' &= D_t[\|\mathbf{v}\|]\mathbf{T} + \|\mathbf{v}\|\mathbf{T}' && \text{Product Rule}\\
&= D_t[\|\mathbf{v}\|]\mathbf{T} + \|\mathbf{v}\|\mathbf{T}'\left(\frac{\|\mathbf{T}'\|}{\|\mathbf{T}'\|}\right)\\
&= D_t[\|\mathbf{v}\|]\mathbf{T} + \|\mathbf{v}\|\,\|\mathbf{T}'\|\,\mathbf{N}. && \mathbf{N} = \mathbf{T}'/\|\mathbf{T}'\|
\end{aligned}$$

Because $\mathbf{a}$ is written as a linear combination of $\mathbf{T}$ and $\mathbf{N}$, it follows that $\mathbf{a}$ lies in the plane determined by $\mathbf{T}$ and $\mathbf{N}$.

The coefficients of $\mathbf{T}$ and $\mathbf{N}$ in the proof of Theorem 12.4 are called the **tangential and normal components of acceleration** and are denoted by $a_{\mathbf{T}} = D_t[\|\mathbf{v}\|]$ and $a_{\mathbf{N}} = \|\mathbf{v}\|\,\|\mathbf{T}'\|$. So, you can write

$$\mathbf{a}(t) = a_{\mathbf{T}}\mathbf{T}(t) + a_{\mathbf{N}}\mathbf{N}(t).$$

The following theorem gives some convenient formulas for $a_{\mathbf{N}}$ and $a_{\mathbf{T}}$.

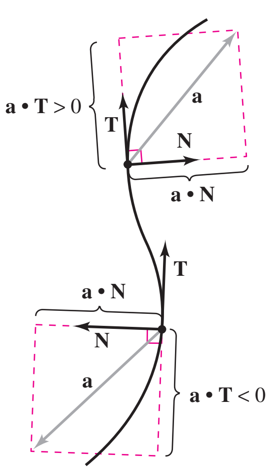
THEOREM 12.5 Tangential and Normal Components of Acceleration

If $\mathbf{r}(t)$ is the position vector for a smooth curve C [for which $\mathbf{N}(t)$ exists], then the tangential and normal components of acceleration are as follows.

$$a_{\mathbf{T}} = D_t[\|\mathbf{v}\|] = \mathbf{a}\cdot\mathbf{T} = \frac{\mathbf{v}\cdot\mathbf{a}}{\|\mathbf{v}\|}$$

$$a_{\mathbf{N}} = \|\mathbf{v}\|\,\|\mathbf{T}'\| = \mathbf{a}\cdot\mathbf{N} = \frac{\|\mathbf{v}\times\mathbf{a}\|}{\|\mathbf{v}\|} = \sqrt{\|\mathbf{a}\|^2 - a_{\mathbf{T}}^2}$$

Note that $a_{\mathbf{N}} \geq 0$. The normal component of acceleration is also called the **centripetal component of acceleration.**



The tangential and normal components of acceleration are obtained by projecting **a** onto **T** and **N**.
Figure 12.24

Proof Note that a lies in the plane of $\mathbf{T}$ and $\mathbf{N}$. So, you can use Figure 12.24 to conclude that, for any time t, the component of the projection of the acceleration vector onto $\mathbf{T}$ is given by $a_{\mathbf{T}} = \mathbf{a}\cdot\mathbf{T}$, and onto $\mathbf{N}$ is given by $a_{\mathbf{N}} = \mathbf{a}\cdot\mathbf{N}$. Moreover, because $\mathbf{a} = \mathbf{v}'$ and $\mathbf{T} = \mathbf{v}/\|\mathbf{v}\|$, you have

$$\begin{aligned}
a_{\mathbf{T}} &= \mathbf{a}\cdot\mathbf{T}\\
&= \mathbf{T}\cdot\mathbf{a}\\
&= \frac{\mathbf{v}}{\|\mathbf{v}\|}\cdot\mathbf{a}\\
&= \frac{\mathbf{v}\cdot\mathbf{a}}{\|\mathbf{v}\|}.
\end{aligned}$$

In Exercises 88 and 89, you are asked to prove the other parts of the theorem.

NOTE The formulas from Theorem 12.5, together with several other formulas from this chapter, are summarized on page 875.

EXAMPLE 5 Tangential and Normal Components of Acceleration

Find the tangential and normal components of acceleration for the position vector given by $\mathbf{r}(t) = 3t\mathbf{i} - t\mathbf{j} + t^2\mathbf{k}$.

Solution Begin by finding the velocity, speed, and acceleration.

$$\mathbf{v}(t) = \mathbf{r}'(t) = 3\mathbf{i} - \mathbf{j} + 2t\mathbf{k}$$
$$\|\mathbf{v}(t)\| = \sqrt{9 + 1 + 4t^2} = \sqrt{10 + 4t^2}$$
$$\mathbf{a}(t) = \mathbf{r}''(t) = 2\mathbf{k}$$

By Theorem 12.5, the tangential component of acceleration is

$$a_{\mathbf{T}} = \frac{\mathbf{v} \cdot \mathbf{a}}{\|\mathbf{v}\|} = \frac{4t}{\sqrt{10 + 4t^2}}$$ Tangential component of acceleration

and because

$$\mathbf{v} \times \mathbf{a} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 3 & -1 & 2t \\ 0 & 0 & 2 \end{vmatrix} = -2\mathbf{i} - 6\mathbf{j}$$

the normal component of acceleration is

$$a_{\mathbf{N}} = \frac{\|\mathbf{v} \times \mathbf{a}\|}{\|\mathbf{v}\|} = \frac{\sqrt{4 + 36}}{\sqrt{10 + 4t^2}} = \frac{2\sqrt{10}}{\sqrt{10 + 4t^2}}.$$ Normal component of acceleration

Try It | Exploration A | Exploration B | Open Exploration

NOTE In Example 5, you could have used the alternative formula for $a_{\mathbf{N}}$ as follows.

$$a_{\mathbf{N}} = \sqrt{\|\mathbf{a}\|^2 - a_{\mathbf{T}}^2} = \sqrt{(2)^2 - \frac{16t^2}{10 + 4t^2}} = \frac{2\sqrt{10}}{\sqrt{10 + 4t^2}}$$

EXAMPLE 6 Finding $a_{\mathbf{T}}$ and $a_{\mathbf{N}}$ for a Circular Helix

Find the tangential and normal components of acceleration for the helix given by $\mathbf{r}(t) = b\cos t\mathbf{i} + b\sin t\mathbf{j} + ct\mathbf{k}$, $b > 0$.

Solution

$$\mathbf{v}(t) = \mathbf{r}'(t) = -b\sin t\mathbf{i} + b\cos t\mathbf{j} + c\mathbf{k}$$
$$\|\mathbf{v}(t)\| = \sqrt{b^2\sin^2 t + b^2\cos^2 t + c^2} = \sqrt{b^2 + c^2}$$
$$\mathbf{a}(t) = \mathbf{r}''(t) = -b\cos t\mathbf{i} - b\sin t\mathbf{j}$$

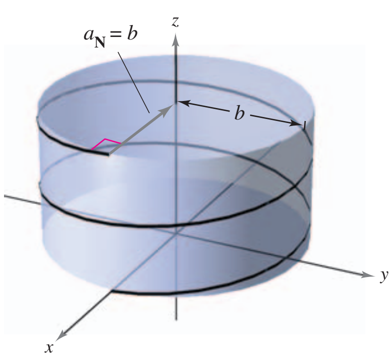
By Theorem 12.5, the tangential component of acceleration is

$$a_{\mathbf{T}} = \frac{\mathbf{v} \cdot \mathbf{a}}{\|\mathbf{v}\|} = \frac{b^2\sin t\cos t - b^2\sin t\cos t + 0}{\sqrt{b^2 + c^2}} = 0.$$ Tangential component of acceleration

Moreover, because $\|\mathbf{a}\| = \sqrt{b^2\cos^2 t + b^2\sin^2 t} = b$, you can use the alternative formula for the normal component of acceleration to obtain

$$a_{\mathbf{N}} = \sqrt{\|\mathbf{a}\|^2 - a_{\mathbf{T}}^2} = \sqrt{b^2 - 0^2} = b.$$ Normal component of acceleration

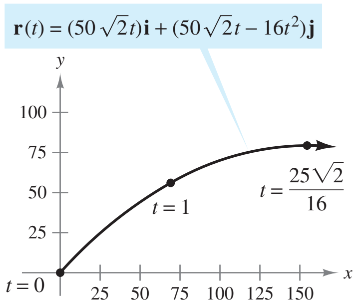
Note that the normal component of acceleration is equal to the magnitude of the acceleration. In other words, because the speed is constant, the acceleration is perpendicular to the velocity. See Figure 12.25.



The normal component of acceleration is equal to the radius of the cylinder around which the helix is spiraling.
Figure 12.25

Rotatable Graph

Try It | Exploration A



The path of a projectile
Figure 12.26

Editable Graph

EXAMPLE 7 Projectile Motion

The position vector for the projectile shown in Figure 12.26 is given by

$$\mathbf{r}(t) = \left(50\sqrt{2}\,t\right)\mathbf{i} + \left(50\sqrt{2}\,t - 16t^2\right)\mathbf{j}. \qquad \text{Position vector}$$

Find the tangential component of acceleration when $t = 0$, 1, and $25\sqrt{2}/16$.

Solution

$$\mathbf{v}(t) = 50\sqrt{2}\,\mathbf{i} + \left(50\sqrt{2} - 32t\right)\mathbf{j} \qquad \text{Velocity vector}$$

$$\|\mathbf{v}(t)\| = 2\sqrt{50^2 - 16(50)\sqrt{2}\,t + 16^2t^2} \qquad \text{Speed}$$

$$\mathbf{a}(t) = -32\mathbf{j} \qquad \text{Acceleration vector}$$

The tangential component of acceleration is

$$a_{\mathbf{T}}(t) = \frac{\mathbf{v}(t)\cdot\mathbf{a}(t)}{\|\mathbf{v}(t)\|} = \frac{-32(50\sqrt{2} - 32t)}{2\sqrt{50^2 - 16(50)\sqrt{2}\,t + 16^2t^2}}. \qquad \text{Tangential component of acceleration}$$

At the specified times, you have

$$a_{\mathbf{T}}(0) = \frac{-32\left(50\sqrt{2}\right)}{100} = -16\sqrt{2} \approx -22.6$$

$$a_{\mathbf{T}}(1) = \frac{-32\left(50\sqrt{2} - 32\right)}{2\sqrt{50^2 - 16(50)\sqrt{2} + 16^2}} \approx -15.4$$

$$a_{\mathbf{T}}\left(\frac{25\sqrt{2}}{16}\right) = \frac{-32\left(50\sqrt{2} - 50\sqrt{2}\right)}{50\sqrt{2}} = 0.$$

You can see from Figure 12.26 that, at the maximum height, when $t = 25\sqrt{2}/16$, the tangential component is 0. This is reasonable because the direction of motion is horizontal at the point and the tangential component of the acceleration is equal to the horizontal component of the acceleration.

Try It

Exploration A

Section 12.5 Arc Length and Curvature

- **Find the arc length of a space curve.**
- **Use the arc length parameter to describe a plane curve or space curve.**
- **Find the curvature of a curve at a point on the curve.**
- **Use a vector-valued function to find frictional force.**

Arc Length

In Section 10.3, you saw that the arc length of a smooth *plane* curve C given by the parametric equations $x = x(t)$ and $y = y(t)$, $a \le t \le b$, is

$$s = \int_a^b \sqrt{[x'(t)]^2 + [y'(t)]^2}\, dt.$$

In vector form, where C is given by $\mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j}$, you can rewrite this equation for arc length as

$$s = \int_a^b \|\mathbf{r}'(t)\|\, dt.$$

The formula for the arc length of a plane curve has a natural extension to a smooth curve in *space*, as stated in the following theorem.

THEOREM 12.6 Arc Length of a Space Curve

If C is a smooth curve given by $\mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j} + z(t)\mathbf{k}$, on an interval $[a, b]$, then the arc length of C on the interval is

$$s = \int_a^b \sqrt{[x'(t)]^2 + [y'(t)]^2 + [z'(t)]^2}\, dt = \int_a^b \|\mathbf{r}'(t)\|\, dt.$$

EXPLORATION

Arc Length Formula The formula for the arc length of a space curve is given in terms of the parametric equations used to represent the curve. Does this mean that the arc length of the curve depends on the parameter being used? Would you want this to be true? Explain your reasoning.

Here is a different parametric representation of the curve in Example 1.

$$\mathbf{r}(t) = t^2\mathbf{i} + \frac{4}{3}t^3\mathbf{j} + \frac{1}{2}t^4\mathbf{k}$$

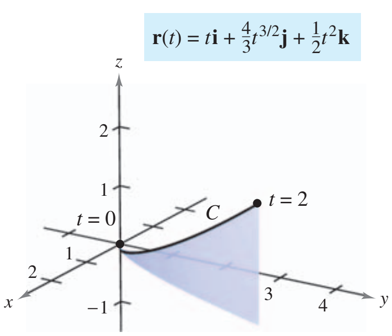
Find the arc length from $t = 0$ to $t = \sqrt{2}$ and compare the result with that found in Example 1.

EXAMPLE 1 Finding the Arc Length of a Curve in Space

Find the arc length of the curve given by

$$\mathbf{r}(t) = t\mathbf{i} + \frac{4}{3}t^{3/2}\mathbf{j} + \frac{1}{2}t^2\mathbf{k}$$

from $t = 0$ to $t = 2$, as shown in Figure 12.27.



As t increases from 0 to 2, the vector $\mathbf{r}(t)$ traces out a curve.
Figure 12.27

Rotatable Graph

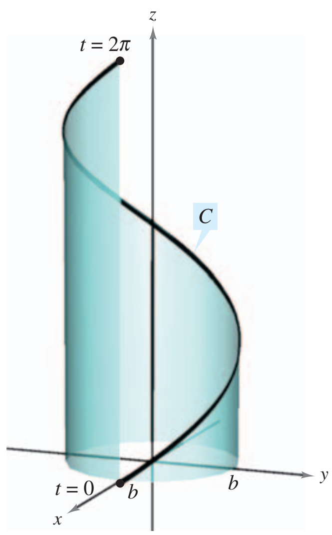
Solution Using $x(t) = t$, $y(t) = \frac{4}{3}t^{3/2}$, and $z(t) = \frac{1}{2}t^2$, you obtain $x'(t) = 1$, $y'(t) = 2t^{1/2}$, and $z'(t) = t$. So, the arc length from $t = 0$ to $t = 2$ is given by

$$\begin{aligned}
s &= \int_0^2 \sqrt{[x'(t)]^2 + [y'(t)]^2 + [z'(t)]^2}\, dt && \text{Formula for arc length}\\
&= \int_0^2 \sqrt{1 + 4t + t^2}\, dt\\
&= \int_0^2 \sqrt{(t+2)^2 - 3}\, dt && \text{Integration tables (Appendix B), Formula 26}\\
&= \left[\frac{t+2}{2}\sqrt{(t+2)^2 - 3} - \frac{3}{2}\ln\left|(t+2) + \sqrt{(t+2)^2 - 3}\right|\right]_0^2\\
&= 2\sqrt{13} - \frac{3}{2}\ln\left(4 + \sqrt{13}\right) - 1 + \frac{3}{2}\ln 3 \approx 4.816.
\end{aligned}$$

Try It Exploration A Open Exploration

EXAMPLE 2 Finding the Arc Length of a Helix

Curve:
$\mathbf{r}(t) = b\cos t\mathbf{i} + b\sin t\mathbf{j} + \sqrt{1-b^2}\,t\mathbf{k}$



One turn of a helix
Figure 12.28

Rotatable Graph

Find the length of one turn of the helix given by

$$\mathbf{r}(t) = b\cos t\mathbf{i} + b\sin t\mathbf{j} + \sqrt{1-b^2}\,t\mathbf{k}$$

as shown in Figure 12.28.

Solution Begin by finding the derivative.

$$\mathbf{r}'(t) = -b\sin t\mathbf{i} + b\cos t\mathbf{j} + \sqrt{1-b^2}\mathbf{k} \qquad \text{Derivative}$$

Now, using the formula for arc length, you can find the length of one turn of the helix by integrating $\|\mathbf{r}'(t)\|$ from 0 to 2π.

$$\begin{aligned} s &= \int_0^{2\pi} \|\mathbf{r}'(t)\|\,dt & \text{Formula for arc length} \\ &= \int_0^{2\pi} \sqrt{b^2(\sin^2 t + \cos^2 t) + (1-b^2)}\,dt \\ &= \int_0^{2\pi} dt \\ &= t\Big]_0^{2\pi} = 2\pi. \end{aligned}$$

So, the length is 2π units.

Try It Exploration A

Arc Length Parameter

You have seen that curves can be represented by vector-valued functions in different ways, depending on the choice of parameter. For *motion* along a curve, the convenient parameter is time t. However, for studying the *geometric properties* of a curve, the convenient parameter is often arc length s.

$s(t) = \int_a^t \sqrt{[x'(u)]^2 + [y'(u)]^2 + [z'(u)]^2}\,du$

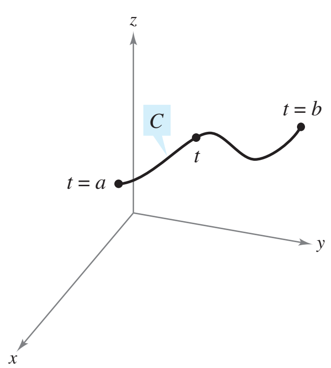


Figure 12.29

Definition of Arc Length Function

Let C be a smooth curve given by $\mathbf{r}(t)$ defined on the closed interval $[a, b]$. For $a \le t \le b$, the **arc length function** is given by

$$s(t) = \int_a^t \|\mathbf{r}'(u)\|\,du = \int_a^t \sqrt{[x'(u)]^2 + [y'(u)]^2 + [z'(u)]^2}\,du.$$

The arc length s is called the **arc length parameter.** (See Figure 12.29.)

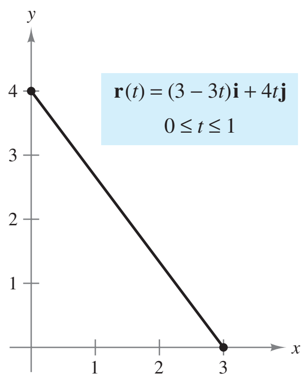
NOTE The arc length function s is *nonnegative*. It measures the distance along C from the initial point $(x(a), y(a), z(a))$ to the point $(x(t), y(t), z(t))$.

Using the definition of the arc length function and the Second Fundamental Theorem of Calculus, you can conclude that

$$\frac{ds}{dt} = \|\mathbf{r}'(t)\|. \qquad \text{Derivative of arc length function}$$

In differential form, you can write

$$ds = \|\mathbf{r}'(t)\|\,dt.$$



The line segment from (3, 0) to (0, 4) can be parametrized using the arc length parameter s.
Figure 12.30

EXAMPLE 3 Finding the Arc Length Function for a Line

Find the arc length function $s(t)$ for the line segment given by

$$\mathbf{r}(t) = (3 - 3t)\mathbf{i} + 4t\mathbf{j}, \quad 0 \le t \le 1$$

and write $\mathbf{r}$ as a function of the parameter s. (See Figure 12.30.)

Solution Because $\mathbf{r}'(t) = -3\mathbf{i} + 4\mathbf{j}$ and

$$\|\mathbf{r}'(t)\| = \sqrt{(-3)^2 + 4^2} = 5$$

you have

$$\begin{aligned} s(t) &= \int_0^t \|\mathbf{r}'(u)\|\, du \\ &= \int_0^t 5\, du \\ &= 5t. \end{aligned}$$

Using $s = 5t$ (or $t = s/5$), you can rewrite $\mathbf{r}$ using the arc length parameter as follows.

$$\mathbf{r}(s) = \left(3 - \tfrac{3}{5}s\right)\mathbf{i} + \tfrac{4}{5}s\,\mathbf{j}, \quad 0 \le s \le 5.$$

Try It **Exploration A**

One of the advantages of writing a vector-valued function in terms of the arc length parameter is that $\|\mathbf{r}'(s)\| = 1$. For instance, in Example 3, you have

$$\|\mathbf{r}'(s)\| = \sqrt{\left(-\frac{3}{5}\right)^2 + \left(\frac{4}{5}\right)^2} = 1.$$

So, for a smooth curve C represented by $\mathbf{r}(s)$, where s is the arc length parameter, the arc length between a and b is

$$\begin{aligned} \text{Length of arc} &= \int_a^b \|\mathbf{r}'(s)\|\, ds \\ &= \int_a^b ds \\ &= b - a \\ &= \text{length of interval.} \end{aligned}$$

Furthermore, if t is *any* parameter such that $\|\mathbf{r}'(t)\| = 1$, then t must be the arc length parameter. These results are summarized in the following theorem, which is stated without proof.

THEOREM 12.7 Arc Length Parameter

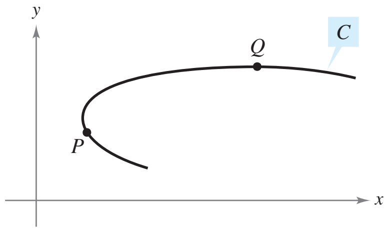
If C is a smooth curve given by

$$\mathbf{r}(s) = x(s)\mathbf{i} + y(s)\mathbf{j} \quad \text{or} \quad \mathbf{r}(s) = x(s)\mathbf{i} + y(s)\mathbf{j} + z(s)\mathbf{k}$$

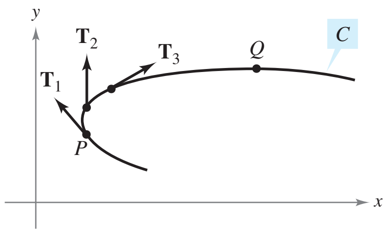
where s is the arc length parameter, then

$$\|\mathbf{r}'(s)\| = 1.$$

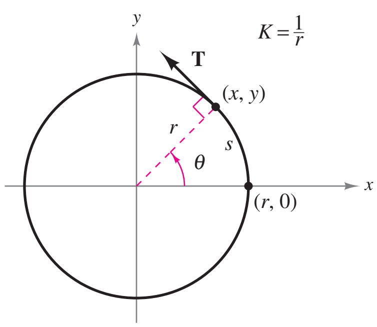
Moreover, if t is *any* parameter for the vector-valued function $\mathbf{r}$ such that $\|\mathbf{r}'(t)\| = 1$, then t must be the arc length parameter.



Curvature at P is greater than at Q.
Figure 12.31



The magnitude of the rate of change of $\mathbf{T}$ with respect to the arc length is the curvature of a curve.
Figure 12.32



The curvature of a circle is constant.
Figure 12.33

Curvature

An important use of the arc length parameter is to find **curvature**—the measure of how sharply a curve bends. For instance, in Figure 12.31 the curve bends more sharply at P than at Q, and you can say that the curvature is greater at P than at Q. You can calculate curvature by calculating the magnitude of the rate of change of the unit tangent vector $\mathbf{T}$ with respect to the arc length s, as shown in Figure 12.32.

Definition of Curvature

Let C be a smooth curve (in the plane *or* in space) given by $\mathbf{r}(s)$, where s is the arc length parameter. The **curvature** K at s is given by

$$K = \left\| \frac{d\mathbf{T}}{ds} \right\| = \|\mathbf{T}'(s)\|.$$

A circle has the same curvature at any point. Moreover, the curvature and the radius of the circle are inversely related. That is, a circle with a large radius has a small curvature, and a circle with a small radius has a large curvature. This inverse relationship is made explicit in the following example.

EXAMPLE 4 Finding the Curvature of a Circle

Show that the curvature of a circle of radius r is $K = 1/r$.

Solution Without loss of generality you can consider the circle to be centered at the origin. Let (x, y) be any point on the circle and let s be the length of the arc from $(r, 0)$ to (x, y), as shown in Figure 12.33. By letting θ be the central angle of the circle, you can represent the circle by

$$\mathbf{r}(\theta) = r\cos\theta\mathbf{i} + r\sin\theta\mathbf{j}. \qquad \theta \text{ is the parameter.}$$

Using the formula for the length of a circular arc $s = r\theta$, you can rewrite $\mathbf{r}(\theta)$ in terms of the arc length parameter as follows.

$$\mathbf{r}(s) = r\cos\frac{s}{r}\mathbf{i} + r\sin\frac{s}{r}\mathbf{j} \qquad \text{Arc length } s \text{ is the parameter.}$$

So, $\mathbf{r}'(s) = -\sin\frac{s}{r}\mathbf{i} + \cos\frac{s}{r}\mathbf{j}$, and it follows that $\|\mathbf{r}'(s)\| = 1$, which implies that the unit tangent vector is

$$\mathbf{T}(s) = \frac{\mathbf{r}'(s)}{\|\mathbf{r}'(s)\|} = -\sin\frac{s}{r}\mathbf{i} + \cos\frac{s}{r}\mathbf{j}$$

and the curvature is given by

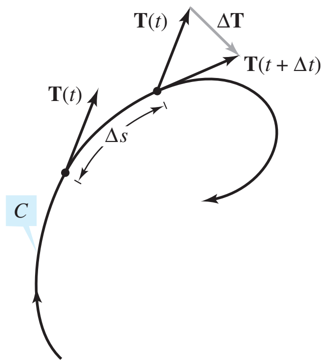
$$K = \|\mathbf{T}'(s)\| = \left\| -\frac{1}{r}\cos\frac{s}{r}\mathbf{i} - \frac{1}{r}\sin\frac{s}{r}\mathbf{j} \right\| = \frac{1}{r}$$

at every point on the circle.

Exploration A Exploration B

NOTE Because a straight line doesn't curve, you would expect its curvature to be 0. Try checking this by finding the curvature of the line given by

$$\mathbf{r}(s) = \left(3 - \frac{3}{5}s\right)\mathbf{i} + \frac{4}{5}s\mathbf{j}.$$



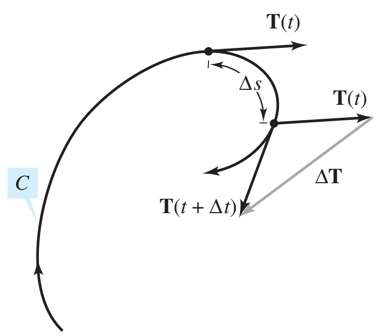


Figure 12.34

In Example 4, the curvature was found by applying the definition directly. This requires that the curve be written in terms of the arc length parameter s. The following theorem gives two other formulas for finding the curvature of a curve written in terms of an arbitrary parameter t. The proof of this theorem is left as an exercise [see Exercise 88, parts (a) and (b)].

THEOREM 12.8 Formulas for Curvature

If C is a smooth curve given by $\mathbf{r}(t)$, then the curvature K of C at t is given by

$$K = \frac{\|\mathbf{T}'(t)\|}{\|\mathbf{r}'(t)\|} = \frac{\|\mathbf{r}'(t) \times \mathbf{r}''(t)\|}{\|\mathbf{r}'(t)\|^3}.$$

Because $\|\mathbf{r}'(t)\| = ds/dt$, the first formula implies that curvature is the ratio of the rate of change in the tangent vector $\mathbf{T}$ to the rate of change in arc length. To see that this is reasonable, let Δt be a "small number." Then,

$$\frac{\mathbf{T}'(t)}{ds/dt} \approx \frac{[\mathbf{T}(t + \Delta t) - \mathbf{T}(t)]/\Delta t}{[s(t + \Delta t) - s(t)]/\Delta t} = \frac{\mathbf{T}(t + \Delta t) - \mathbf{T}(t)}{s(t + \Delta t) - s(t)} = \frac{\Delta \mathbf{T}}{\Delta s}.$$

In other words, for a given Δs, the greater the length of $\Delta \mathbf{T}$, the more the curve bends at t, as shown in Figure 12.34.

EXAMPLE 5 Finding the Curvature of a Space Curve

Find the curvature of the curve given by $\mathbf{r}(t) = 2t\mathbf{i} + t^2\mathbf{j} - \frac{1}{3}t^3\mathbf{k}$.

Solution It is not apparent whether this parameter represents arc length, so you should use the formula $K = \|\mathbf{T}'(t)\|/\|\mathbf{r}'(t)\|$.

$$\mathbf{r}'(t) = 2\mathbf{i} + 2t\mathbf{j} - t^2\mathbf{k}$$

$$\|\mathbf{r}'(t)\| = \sqrt{4 + 4t^2 + t^4} = t^2 + 2 \qquad \text{Length of } \mathbf{r}'(t)$$

$$\mathbf{T}(t) = \frac{\mathbf{r}'(t)}{\|\mathbf{r}'(t)\|} = \frac{2\mathbf{i} + 2t\mathbf{j} - t^2\mathbf{k}}{t^2 + 2}$$

$$\mathbf{T}'(t) = \frac{(t^2 + 2)(2\mathbf{j} - 2t\mathbf{k}) - (2t)(2\mathbf{i} + 2t\mathbf{j} - t^2\mathbf{k})}{(t^2 + 2)^2}$$

$$= \frac{-4t\mathbf{i} + (4 - 2t^2)\mathbf{j} - 4t\mathbf{k}}{(t^2 + 2)^2}$$

$$\|\mathbf{T}'(t)\| = \frac{\sqrt{16t^2 + 16 - 16t^2 + 4t^4 + 16t^2}}{(t^2 + 2)^2}$$

$$= \frac{2(t^2 + 2)}{(t^2 + 2)^2}$$

$$= \frac{2}{t^2 + 2} \qquad \text{Length of } \mathbf{T}'(t)$$

Therefore,

$$K = \frac{\|\mathbf{T}'(t)\|}{\|\mathbf{r}'(t)\|} = \frac{2}{(t^2 + 2)^2}. \qquad \text{Curvature}$$

Try It **Exploration A**

The following theorem presents a formula for calculating the curvature of a plane curve given by $y = f(x)$.

THEOREM 12.9 Curvature in Rectangular Coordinates

If C is the graph of a twice-differentiable function given by $y = f(x)$, then the curvature K at the point (x, y) is given by

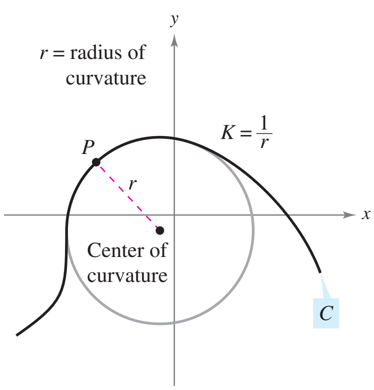
$$K = \frac{|y''|}{[1 + (y')^2]^{3/2}}.$$

Proof By representing the curve C by $\mathbf{r}(x) = x\mathbf{i} + f(x)\mathbf{j} + 0\mathbf{k}$ (where x is the parameter), you obtain $\mathbf{r}'(x) = \mathbf{i} + f'(x)\mathbf{j}$,

$$\|\mathbf{r}'(x)\| = \sqrt{1 + [f'(x)]^2}$$

and $\mathbf{r}''(x) = f''(x)\mathbf{j}$. Because $\mathbf{r}'(x) \times \mathbf{r}''(x) = f''(x)\mathbf{k}$, it follows that the curvature is

$$\begin{aligned} K &= \frac{\|\mathbf{r}'(x) \times \mathbf{r}''(x)\|}{\|\mathbf{r}'(x)\|^3} \\ &= \frac{|f''(x)|}{\{1 + [f'(x)]^2\}^{3/2}} \\ &= \frac{|y''|}{[1 + (y')^2]^{3/2}}. \end{aligned}$$



The circle of curvature
Figure 12.35

Let C be a curve with curvature K at point P. The circle passing through point P with radius $r = 1/K$ is called the **circle of curvature** if the circle lies on the concave side of the curve and shares a common tangent line with the curve at point P. The radius is called the **radius of curvature** at P, and the center of the circle is called the **center of curvature.**

The circle of curvature gives you a nice way to estimate graphically the curvature K at a point P on a curve. Using a compass, you can sketch a circle that lies against the concave side of the curve at point P, as shown in Figure 12.35. If the circle has a radius of r, you can estimate the curvature to be $K = 1/r$.

EXAMPLE 6 Finding Curvature in Rectangular Coordinates

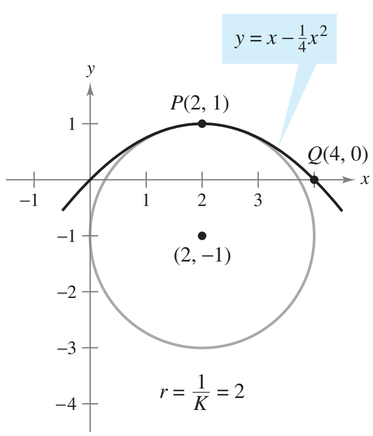
Find the curvature of the parabola given by $y = x - \frac{1}{4}x^2$ at $x = 2$. Sketch the circle of curvature at $(2, 1)$.

Solution The curvature at $x = 2$ is as follows.

$$y' = 1 - \frac{x}{2} \qquad\qquad y' = 0$$

$$y'' = -\frac{1}{2} \qquad\qquad y'' = -\frac{1}{2}$$

$$K = \frac{|y''|}{[1 + (y')^2]^{3/2}} \qquad\qquad K = \frac{1}{2}$$

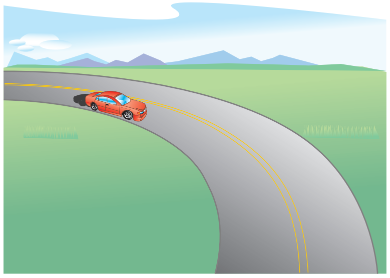


The circle of curvature
Figure 12.36

Because the curvature at $P(2, 1)$ is $\frac{1}{2}$, it follows that the radius of the circle of curvature at that point is 2. So, the center of curvature is $(2, -1)$, as shown in Figure 12.36. [In the figure, note that the curve has the greatest curvature at P. Try showing that the curvature at $Q(4, 0)$ is $1/2^{5/2} \approx 0.177$.]

Editable Graph

Try It Exploration A Exploration B



The amount of thrust felt by passengers in a car that is turning depends on two things—the speed of the car and the sharpness of the turn.
Figure 12.37

Arc length and curvature are closely related to the tangential and normal components of acceleration. The tangential component of acceleration is the rate of change of the speed, which in turn is the rate of change of the arc length. This component is negative as a moving object slows down and positive as it speeds up—regardless of whether the object is turning or traveling in a straight line. So, the tangential component is solely a function of the arc length and is independent of the curvature.

On the other hand, the normal component of acceleration is a function of *both* speed and curvature. This component measures the acceleration acting perpendicular to the direction of motion. To see why the normal component is affected by both speed and curvature, imagine that you are driving a car around a turn, as shown in Figure 12.37. If your speed is high and the turn is sharp, you feel yourself thrown against the car door. By lowering your speed *or* taking a more gentle turn, you are able to lessen this sideways thrust.

The next theorem explicitly states the relationships among speed, curvature, and the components of acceleration.

NOTE Note that Theorem 12.10 gives additional formulas for $a_{\mathbf{T}}$ and $a_{\mathbf{N}}$.

THEOREM 12.10 Acceleration, Speed, and Curvature

If $\mathbf{r}(t)$ is the position vector for a smooth curve C, then the acceleration vector is given by

$$\mathbf{a}(t) = \frac{d^2s}{dt^2}\mathbf{T} + K\left(\frac{ds}{dt}\right)^2\mathbf{N}$$

where K is the curvature of C and ds/dt is the speed.

Proof For the position vector $\mathbf{r}(t)$, you have

$$\begin{aligned}\mathbf{a}(t) &= a_{\mathbf{T}}\mathbf{T} + a_{\mathbf{N}}\mathbf{N}\\ &= D_t[\|\mathbf{v}\|]\mathbf{T} + \|\mathbf{v}\|\,\|\mathbf{T}'\|\mathbf{N}\\ &= \frac{d^2s}{dt^2}\mathbf{T} + \frac{ds}{dt}(\|\mathbf{v}\|K)\mathbf{N}\\ &= \frac{d^2s}{dt^2}\mathbf{T} + K\left(\frac{ds}{dt}\right)^2\mathbf{N}.\end{aligned}$$

EXAMPLE 7 Tangential and Normal Components of Acceleration

Find $a_{\mathbf{T}}$ and $a_{\mathbf{N}}$ for the curve given by

$$\mathbf{r}(t) = 2t\mathbf{i} + t^2\mathbf{j} - \tfrac{1}{3}t^3\mathbf{k}.$$

Solution From Example 5, you know that

$$\frac{ds}{dt} = \|\mathbf{r}'(t)\| = t^2 + 2 \quad \text{and} \quad K = \frac{2}{(t^2+2)^2}.$$

Therefore,

$$a_{\mathbf{T}} = \frac{d^2s}{dt^2} = 2t \qquad \text{Tangential component}$$

and

$$a_{\mathbf{N}} = K\left(\frac{ds}{dt}\right)^2 = \frac{2}{(t^2+2)^2}(t^2+2)^2 = 2. \qquad \text{Normal component}$$

Try It Exploration A

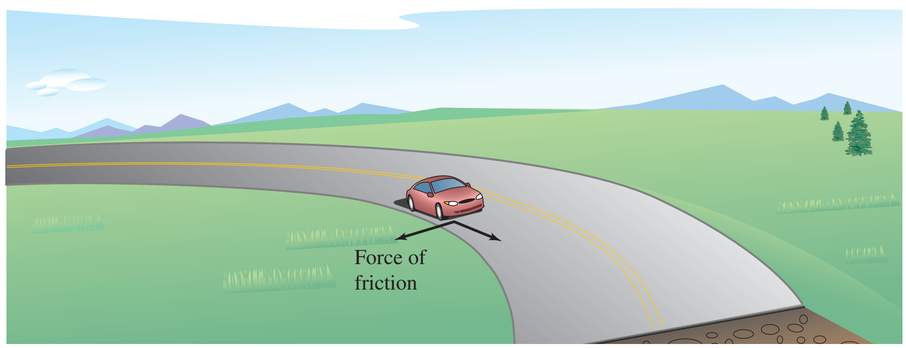
Application

There are many applications in physics and engineering dynamics that involve the relationships among speed, arc length, curvature, and acceleration. One such application concerns frictional force.

A moving object with mass m is in contact with a stationary object. The total force required to produce an acceleration $\mathbf{a}$ along a given path is

$$\begin{aligned}\mathbf{F} = m\mathbf{a} &= m\left(\frac{d^2s}{dt^2}\right)\mathbf{T} + mK\left(\frac{ds}{dt}\right)^2\mathbf{N} \\ &= ma_{\mathbf{T}}\mathbf{T} + ma_{\mathbf{N}}\mathbf{N}.\end{aligned}$$

The portion of this total force that is supplied by the stationary object is called the **force of friction.** For example, if a car moving with constant speed is rounding a turn, the roadway exerts a frictional force that keeps the car from sliding off the road. If the car is not sliding, the frictional force is perpendicular to the direction of motion and has magnitude equal to the normal component of acceleration, as shown in Figure 12.38. The potential frictional force of a road around a turn can be increased by banking the roadway.



The force of friction is perpendicular to the direction of the motion.
Figure 12.38

EXAMPLE 8 Frictional Force

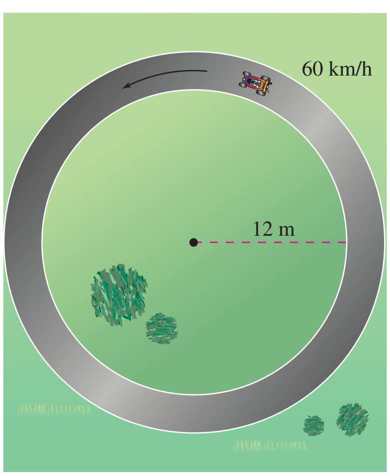


Figure 12.39

A 360-kilogram go-cart is driven at a speed of 60 kilometers per hour around a circular racetrack of radius 12 meters, as shown in Figure 12.39. To keep the cart from skidding off course, what frictional force must the track surface exert on the tires?

Solution The frictional force must equal the mass times the normal component of acceleration. For this circular path, you know that the curvature is

$$K = \frac{1}{12}. \qquad \text{Curvature of circular racetrack}$$

Therefore, the frictional force is

$$\begin{aligned}ma_{\mathbf{N}} &= mK\left(\frac{ds}{dt}\right)^2 \\ &= (360 \text{ kg})\left(\frac{1}{12 \text{ m}}\right)\left(\frac{60{,}000 \text{ m}}{3600 \text{ sec}}\right)^2 \\ &\approx 8333 \text{ (kg)(m)/sec}^2.\end{aligned}$$

Try It Exploration A

Summary of Velocity, Acceleration, and Curvature

Let C be a curve (in the plane or in space) given by the position function

$\mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j}$ Curve in the plane

$\mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j} + z(t)\mathbf{k}.$ Curve in space

Velocity vector, speed, and acceleration vector:

$\mathbf{v}(t) = \mathbf{r}'(t)$ Velocity vector

$\|\mathbf{v}(t)\| = \dfrac{ds}{dt} = \|\mathbf{r}'(t)\|$ Speed

$\mathbf{a}(t) = \mathbf{r}''(t) = a_{\mathbf{T}}\mathbf{T}(t) + a_{\mathbf{N}}\mathbf{N}(t)$ Acceleration vector

Unit tangent vector and principal unit normal vector:

$$\mathbf{T}(t) = \frac{\mathbf{r}'(t)}{\|\mathbf{r}'(t)\|} \quad \text{and} \quad \mathbf{N}(t) = \frac{\mathbf{T}'(t)}{\|\mathbf{T}'(t)\|}$$

Components of acceleration:

$$a_{\mathbf{T}} = \mathbf{a} \cdot \mathbf{T} = \frac{\mathbf{v} \cdot \mathbf{a}}{\|\mathbf{v}\|} = \frac{d^2s}{dt^2}$$

$$a_{\mathbf{N}} = \mathbf{a} \cdot \mathbf{N} = \frac{\|\mathbf{v} \times \mathbf{a}\|}{\|\mathbf{v}\|} = \sqrt{\|\mathbf{a}\|^2 - a_{\mathbf{T}}^{\,2}} = K\left(\frac{ds}{dt}\right)^2$$

Formulas for curvature in the plane:

$K = \dfrac{|y''|}{[1 + (y')^2]^{3/2}}$ C given by $y = f(x)$

$K = \dfrac{|x'y'' - y'x''|}{[(x')^2 + (y')^2]^{3/2}}$ C given by $x = x(t)$, $y = y(t)$

Formulas for curvature in the plane or in space:

$K = \|\mathbf{T}'(s)\| = \|\mathbf{r}''(s)\|$ s is arc length parameter.

$K = \dfrac{\|\mathbf{T}'(t)\|}{\|\mathbf{r}'(t)\|} = \dfrac{\|\mathbf{r}'(t) \times \mathbf{r}''(t)\|}{\|\mathbf{r}'(t)\|^3}$ t is general parameter.

$$K = \frac{\mathbf{a}(t) \cdot \mathbf{N}(t)}{\|\mathbf{v}(t)\|^2}$$

Cross product formulas apply only to curves in space.

Section 13.1 Introduction to Functions of Several Variables

- **Understand the notation for a function of several variables.**
- **Sketch the graph of a function of two variables.**
- **Sketch level curves for a function of two variables.**
- **Sketch level surfaces for a function of three variables.**
- **Use computer graphics to graph a function of two variables.**

EXPLORATION

Comparing Dimensions Without using a graphing utility, describe the graph of each function of two variables.

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

b. $z = x + y$

c. $z = x^2 + y$

d. $z = \sqrt{x^2 + y^2}$

e. $z = \sqrt{1 - x^2 + y^2}$

Functions of Several Variables

So far in this text, you have dealt only with functions of a single (independent) variable. Many familiar quantities, however, are functions of two or more variables. For instance, the work done by a force $(W = FD)$ and the volume of a right circular cylinder $(V = \pi r^2 h)$ are both functions of two variables. The volume of a rectangular solid $(V = lwh)$ is a function of three variables. The notation for a function of two or more variables is similar to that for a function of a single variable. Here are two examples.

$$z = f\underbrace{(x, y)}_{\text{2 variables}} = x^2 + xy \qquad \text{Function of two variables}$$

and

$$w = f\underbrace{(x, y, z)}_{\text{3 variables}} = x + 2y - 3z \qquad \text{Function of three variables}$$

Definition of a Function of Two Variables

Let D be a set of ordered pairs of real numbers. If to each ordered pair (x, y) in D there corresponds a unique real number $f(x, y)$, then f is called a **function of *x* and *y*.** The set D is the **domain** of f, and the corresponding set of values for $f(x, y)$ is the **range** of f.

MARY FAIRFAX SOMERVILLE (1780–1872)

Somerville was interested in the problem of creating geometric models for functions of several variables. Her most well-known book, *The Mechanics of the Heavens*, was published in 1831.

MathBio

For the function given by $z = f(x, y)$, x and y are called the **independent variables** and z is called the **dependent variable.**

Similar definitions can be given for functions of three, four, or n variables, where the domains consist of ordered triples (x_1, x_2, x_3), quadruples (x_1, x_2, x_3, x_4), and n-tuples $(x_1, x_2, . . . , x_n)$. In all cases, the range is a set of real numbers. In this chapter, you will study only functions of two or three variables.

As with functions of one variable, the most common way to describe a function of several variables is with an *equation*, and unless otherwise restricted, you can assume that the domain is the set of all points for which the equation is defined. For instance, the domain of the function given by

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

is assumed to be the entire xy-plane. Similarly, the domain of

$$f(x, y) = \ln xy$$

is the set of all points (x, y) in the plane for which $xy > 0$. This consists of all points in the first and third quadrants.

EXAMPLE 1 Domains of Functions of Several Variables

Find the domain of each function.

a. $f(x, y) = \dfrac{\sqrt{x^2 + y^2 - 9}}{x}$ **b.** $g(x, y, z) = \dfrac{x}{\sqrt{9 - x^2 - y^2 - z^2}}$

Solution

a. The function f is defined for all points (x, y) such that $x \neq 0$ and

$$x^2 + y^2 \geq 9.$$

So, the domain is the set of all points lying on or outside the circle $x^2 + y^2 = 9$, *except* those points on the y-axis, as shown in Figure 13.1.

b. The function g is defined for all points (x, y, z) such that

$$x^2 + y^2 + z^2 < 9.$$

Consequently, the domain is the set of all points (x, y, z) lying inside a sphere of radius 3 that is centered at the origin.

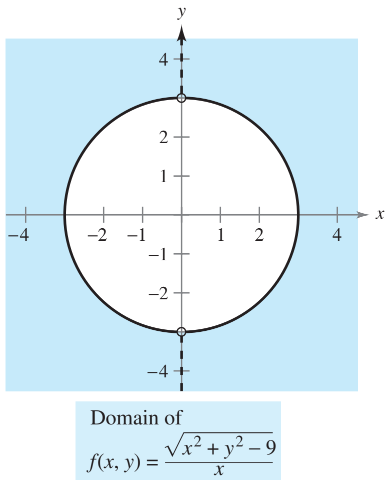


Figure 13.1

Try It Exploration A

Functions of several variables can be combined in the same ways as functions of single variables. For instance, you can form the sum, difference, product, and quotient of two functions of two variables as follows.

$(f \pm g)(x, y) = f(x, y) \pm g(x, y)$		Sum or difference
$(fg)(x, y) = f(x, y)g(x, y)$		Product
$\dfrac{f}{g}(x, y) = \dfrac{f(x, y)}{g(x, y)}$	$g(x, y) \neq 0$	Quotient

You cannot form the composite of two functions of several variables. However, if h is a function of several variables and g is a function of a single variable, you can form the **composite** function $(g \circ h)(x, y)$ as follows.

$$(g \circ h)(x, y) = g(h(x, y)) \qquad \text{Composition}$$

The domain of this composite function consists of all (x, y) in the domain of h such that $h(x, y)$ is in the domain of g. For example, the function given by

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

can be viewed as the composite of the function of two variables given by $h(x, y) = 16 - 4x^2 - y^2$ and the function of a single variable given by $g(u) = \sqrt{u}$. The domain of this function is the set of all points lying on or inside the ellipse given by $4x^2 + y^2 = 16$.

A function that can be written as a sum of functions of the form $cx^m y^n$ (where c is a real number and m and n are nonnegative integers) is called a **polynomial function** of two variables. For instance, the functions given by

$$f(x, y) = x^2 + y^2 - 2xy + x + 2 \quad \text{and} \quad g(x, y) = 3xy^2 + x - 2$$

are polynomial functions of two variables. A **rational function** is the quotient of two polynomial functions. Similar terminology is used for functions of more than two variables.

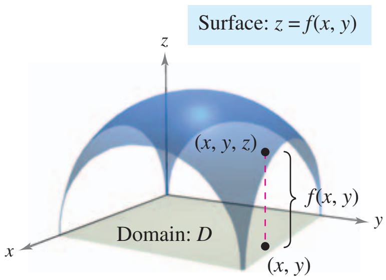


Figure 13.2

Rotatable Graph

The Graph of a Function of Two Variables

As with functions of a single variable, you can learn a lot about the behavior of a function of two variables by sketching its graph. The **graph** of a function f of two variables is the set of all points (x, y, z) for which $z = f(x, y)$ and (x, y) is in the domain of f. This graph can be interpreted geometrically as a *surface in space*, as discussed in Sections 11.5 and 11.6. In Figure 13.2, note that the graph of $z = f(x, y)$ is a surface whose projection onto the xy-plane is D, the domain of f. To each point (x, y) in D there corresponds a point (x, y, z) on the surface, and, conversely, to each point (x, y, z) on the surface there corresponds a point (x, y) in D.

EXAMPLE 2 Describing the Graph of a Function of Two Variables

What is the range of $f(x, y) = \sqrt{16 - 4x^2 - y^2}$? Describe the graph of f.

Solution The domain D implied by the equation for f is the set of all points (x, y) such that $16 - 4x^2 - y^2 \geq 0$. So, D is the set of all points lying on or inside the ellipse given by

$$\frac{x^2}{4} + \frac{y^2}{16} = 1. \qquad \text{Ellipse in the } xy\text{-plane}$$

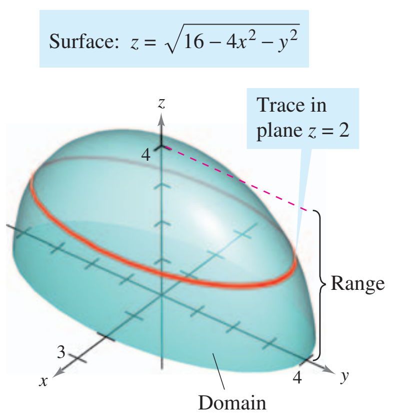
The range of f is all values $z = f(x, y)$ such that $0 \leq z \leq \sqrt{16}$ or

$$0 \leq z \leq 4. \qquad \text{Range of } f$$

A point (x, y, z) is on the graph of f if and only if

$$\begin{aligned} z &= \sqrt{16 - 4x^2 - y^2} \\ z^2 &= 16 - 4x^2 - y^2 \\ 4x^2 + y^2 + z^2 &= 16 \\ \frac{x^2}{4} + \frac{y^2}{16} + \frac{z^2}{16} &= 1, \qquad 0 \leq z \leq 4. \end{aligned}$$

From Section 11.6, you know that the graph of f is the upper half of an ellipsoid, as shown in Figure 13.3.



The graph of $f(x, y) = \sqrt{16 - 4x^2 - y^2}$ is the upper half of an ellipsoid.
Figure 13.3

Rotatable Graph

Try It Exploration A

To sketch a surface in space *by hand,* it helps to use traces in planes parallel to the coordinate planes, as shown in Figure 13.3. For example, to find the trace of the surface in the plane $z = 2$, substitute $z = 2$ in the equation $z = \sqrt{16 - 4x^2 - y^2}$ and obtain

$$2 = \sqrt{16 - 4x^2 - y^2} \quad \Rightarrow \quad \frac{x^2}{3} + \frac{y^2}{12} = 1.$$

So, the trace is an ellipse centered at the point $(0, 0, 2)$ with major and minor axes of lengths $4\sqrt{3}$ and $2\sqrt{3}$.

Traces are also used with most three-dimensional graphing utilities. For instance, Figure 13.4 shows a computer-generated version of the surface given in Example 2. For this graph, the computer took 25 traces parallel to the xy-plane and 12 traces in vertical planes.

If you have access to a three-dimensional graphing utility, use it to graph several surfaces.

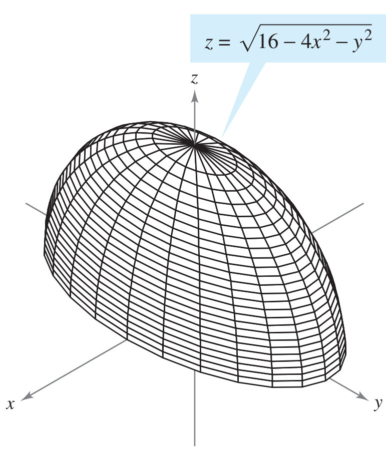
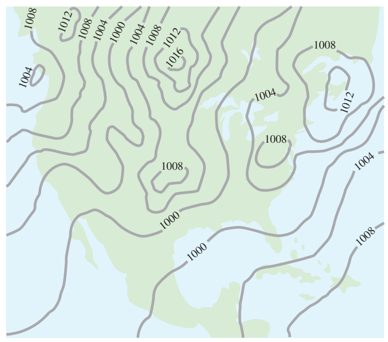


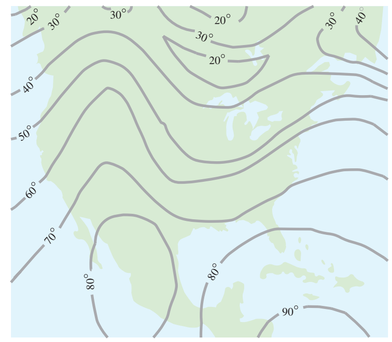
Figure 13.4

Level Curves

A second way to visualize a function of two variables is to use a **scalar field** in which the scalar $z = f(x, y)$ is assigned to the point (x, y). A scalar field can be characterized by **level curves** (or **contour lines**) along which the value of $f(x, y)$ is constant. For instance, the weather map in Figure 13.5 shows level curves of equal pressure called **isobars.** In weather maps for which the level curves represent points of equal temperature, the level curves are called **isotherms,** as shown in Figure 13.6. Another common use of level curves is in representing electric potential fields. In this type of map, the level curves are called **equipotential lines.**



Level curves show the lines of equal pressure (isobars) measured in millibars.
Figure 13.5



Level curves show the lines of equal temperature (isotherms) measured in degrees Fahrenheit.
Figure 13.6

Contour maps are commonly used to show regions on Earth's surface, with the level curves representing the height above sea level. This type of map is called a **topographic map.** For example, the mountain shown in Figure 13.7 is represented by the topographic map in Figure 13.8. View the animation to see this more clearly.

A contour map depicts the variation of z with respect to x and y by the spacing between level curves. Much space between level curves indicates that z is changing slowly, whereas little space indicates a rapid change in z. Furthermore, to give a good three-dimensional illusion in a contour map, it is important to choose c-values that are *evenly spaced*.

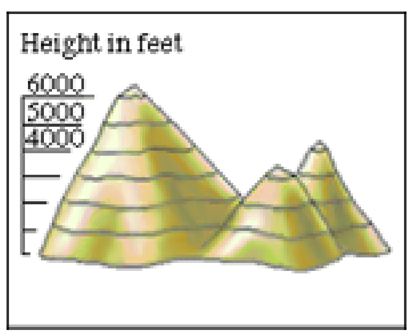


Figure 13.7

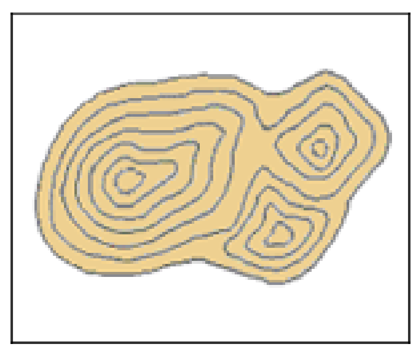


Figure 13.8

Animation

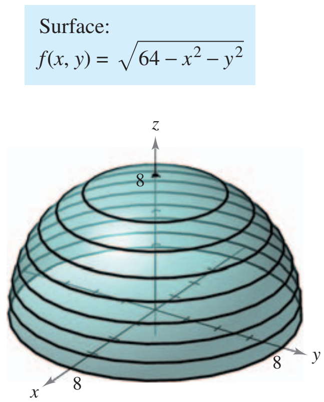
EXAMPLE 3 Sketching a Contour Map

The hemisphere given by $f(x, y) = \sqrt{64 - x^2 - y^2}$ is shown in Figure 13.9. Sketch a contour map for this surface using level curves corresponding to $c = 0, 1, 2, \ldots, 8$.

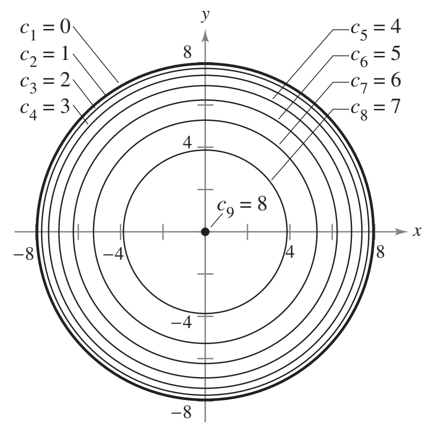
Solution For each value of c, the equation given by $f(x, y) = c$ is a circle (or point) in the xy-plane. For example, when $c_1 = 0$, the level curve is

$$x^2 + y^2 = 64 \qquad \text{Circle of radius 8}$$

which is a circle of radius 8. Figure 13.10 shows the nine level curves for the hemisphere.



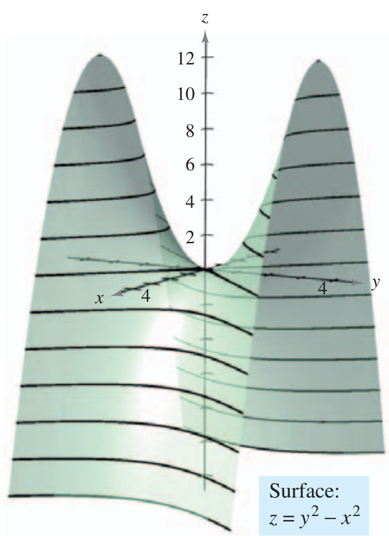
Hemisphere
Figure 13.9



Contour map
Figure 13.10

Animation

Try It Exploration A



Hyperbolic paraboloid
Figure 13.11

Animation

EXAMPLE 4 Sketching a Contour Map

The hyperbolic paraboloid given by

$$z = y^2 - x^2$$

is shown in Figure 13.11. Sketch a contour map for this surface.

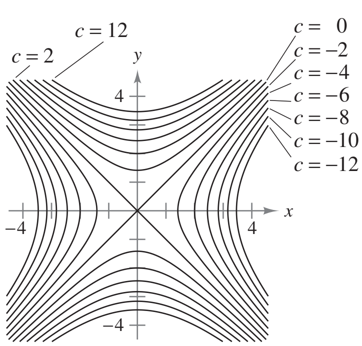
Solution For each value of c, let $f(x, y) = c$ and sketch the resulting level curve in the xy-plane. For this function, each of the level curves $(c \neq 0)$ is a hyperbola whose asymptotes are the lines $y = \pm x$. If $c < 0$, the transverse axis is horizontal. For instance, the level curve for $c = -4$ is given by

$$\frac{x^2}{2^2} - \frac{y^2}{2^2} = 1. \qquad \text{Hyperbola with horizontal transverse axis}$$

If $c > 0$, the transverse axis is vertical. For instance, the level curve for $c = 4$ is given by

$$\frac{y^2}{2^2} - \frac{x^2}{2^2} = 1. \qquad \text{Hyperbola with vertical transverse axis}$$

If $c = 0$, the level curve is the degenerate conic representing the intersecting asymptotes, as shown in Figure 13.12.



Hyperbolic level curves (at increments of 2)
Figure 13.12

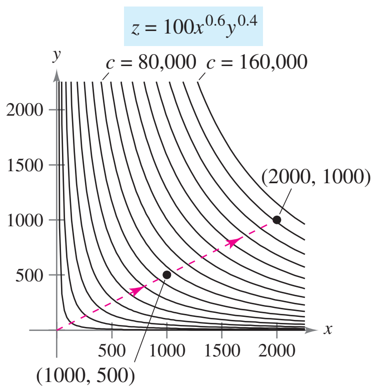
Try It Open Exploration

One example of a function of two variables used in economics is the **Cobb-Douglas production function.** This function is used as a model to represent the number of units produced by varying amounts of labor and capital. If x measures the units of labor and y measures the units of capital, the number of units produced is given by

$$f(x, y) = Cx^a y^{1-a}$$

where C and a are constants with $0 < a < 1$.

EXAMPLE 5 The Cobb-Douglas Production Function



Level curves (at increments of 10,000)
Figure 13.13

A toy manufacturer estimates a production function to be $f(x, y) = 100x^{0.6}y^{0.4}$, where x is the number of units of labor and y is the number of units of capital. Compare the production level when $x = 1000$ and $y = 500$ with the production level when $x = 2000$ and $y = 1000$.

Solution When $x = 1000$ and $y = 500$, the production level is

$$f(1000, 500) = 100(1000^{0.6})(500^{0.4}) \approx 75{,}786.$$

When $x = 2000$ and $y = 1000$, the production level is

$$f(2000, 1000) = 100(2000^{0.6})(1000^{0.4}) = 151{,}572.$$

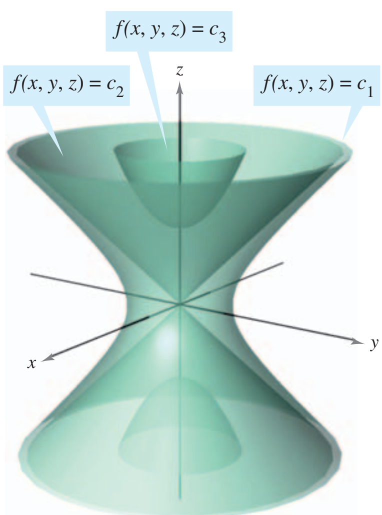
The level curves of $z = f(x, y)$ are shown in Figure 13.13. Note that by doubling both x and y, you double the production level (see Exercise 79).

Try It **Exploration A**

Level Surfaces

The concept of a level curve can be extended by one dimension to define a **level surface.** If f is a function of three variables and c is a constant, the graph of the equation $f(x, y, z) = c$ is a **level surface** of the function f, as shown in Figure 13.14.

With computers, engineers and scientists have developed other ways to view functions of three variables. For instance, Figure 13.15 shows a computer simulation that uses color to represent the pressure waves of a high-speed train traveling through a tunnel.



Level surfaces of f
Figure 13.14

Rotatable Graph

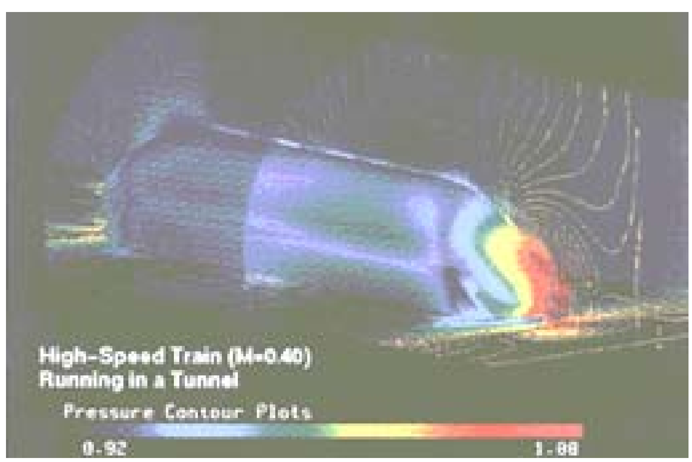


Figure 13.15

EXAMPLE 6 Level Surfaces

Describe the level surfaces of the function

$$f(x, y, z) = 4x^2 + y^2 + z^2.$$

Solution Each level surface has an equation of the form

$$4x^2 + y^2 + z^2 = c. \qquad \text{Equation of level surface}$$

So, the level surfaces are ellipsoids (whose cross sections parallel to the yz-plane are circles). As c increases, the radii of the circular cross sections increase according to the square root of c. For example, the level surfaces corresponding to the values $c = 0$, $c = 4$, and $c = 16$ are as follows.

$$4x^2 + y^2 + z^2 = 0 \qquad \text{Level surface for } c = 0 \text{ (single point)}$$

$$\frac{x^2}{1} + \frac{y^2}{4} + \frac{z^2}{4} = 1 \qquad \text{Level surface for } c = 4 \text{ (ellipsoid)}$$

$$\frac{x^2}{4} + \frac{y^2}{16} + \frac{z^2}{16} = 1 \qquad \text{Level surface for } c = 16 \text{ (ellipsoid)}$$

These level surfaces are shown in Figure 13.16.

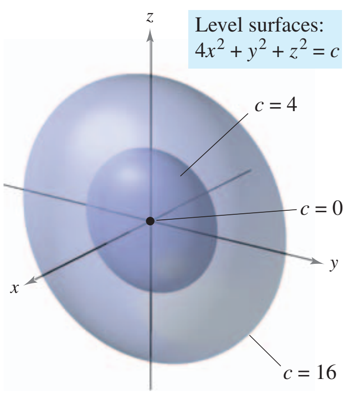


Figure 13.16

Rotatable Graph

Try It

Exploration A

NOTE If the function in Example 6 represented the *temperature* at the point (x, y, z), the level surfaces shown in Figure 13.16 would be called **isothermal surfaces.**

Computer Graphics

The problem of sketching the graph of a surface in space can be simplified by using a computer. Although there are several types of three-dimensional graphing utilities, most use some form of trace analysis to give the illusion of three dimensions. To use such a graphing utility, you usually need to enter the equation of the surface, the region in the xy-plane over which the surface is to be plotted, and the number of traces to be taken. For instance, to graph the surface given by

$$f(x, y) = (x^2 + y^2)e^{1 - x^2 - y^2}$$

you might choose the following bounds for x, y, and z.

$$-3 \le x \le 3 \qquad \text{Bounds for } x$$

$$-3 \le y \le 3 \qquad \text{Bounds for } y$$

$$0 \le z \le 3 \qquad \text{Bounds for } z$$

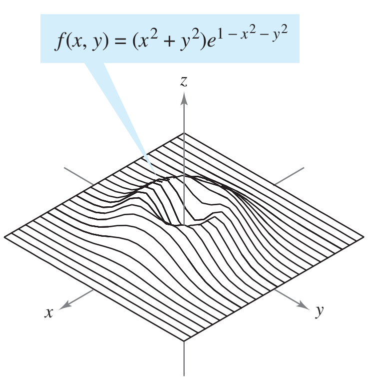
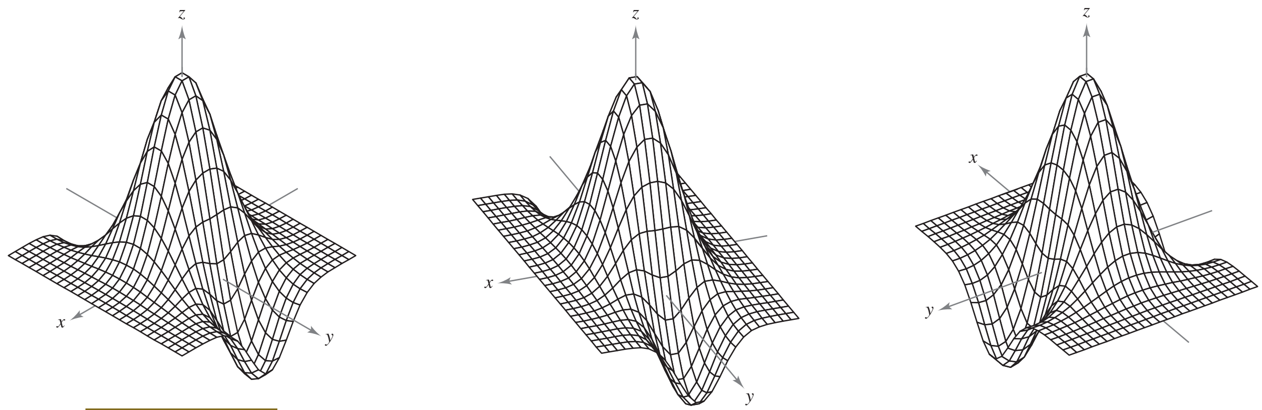


Figure 13.17

Rotatable Graph

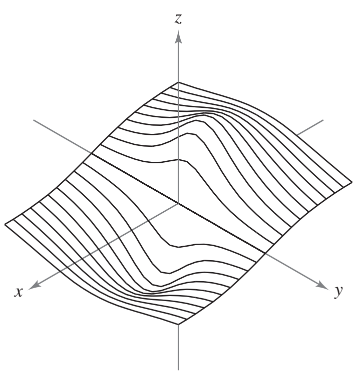
Figure 13.17 shows a computer-generated graph of this surface using 26 traces taken parallel to the yz-plane. To heighten the three-dimensional effect, the program uses a "hidden line" routine. That is, it begins by plotting the traces in the foreground (those corresponding to the largest x-values), and then, as each new trace is plotted, the program determines whether all or only part of the next trace should be shown.

The graphs on page 891 show a variety of surfaces that were plotted by computer. If you have access to a computer drawing program, use it to reproduce these surfaces. Remember also that the three-dimensional graphics in this text can be viewed and rotated.



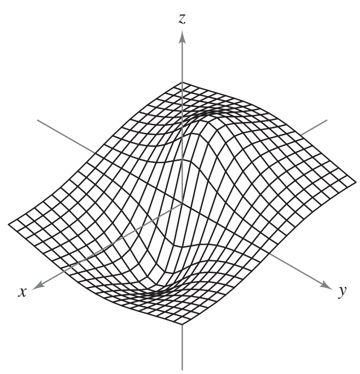
Rotatable Graph

Three different views of the graph of $f(x, y) = (2 - y^2 + x^2)e^{1 - x^2 - (y^2/4)}$



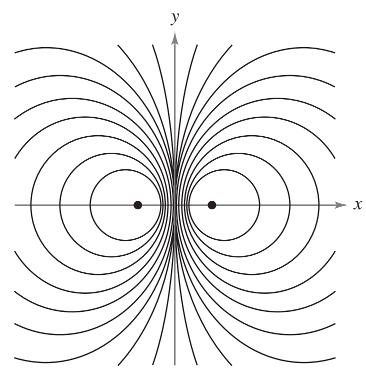
Single traces

Rotatable Graph



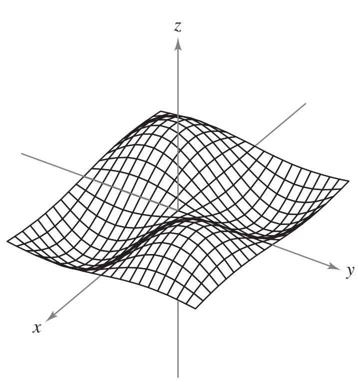
Double traces

Rotatable Graph



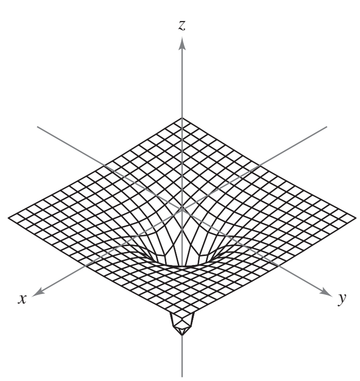
Level curves

Traces and level curves of the graph of $f(x, y) = \dfrac{-4x}{x^2 + y^2 + 1}$



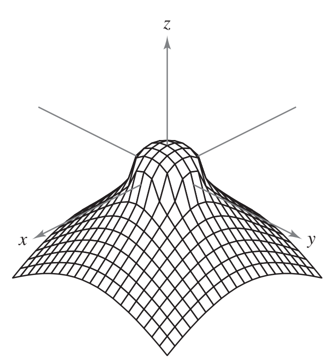
$f(x, y) = \sin x \sin y$

Rotatable Graph



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

Rotatable Graph



$f(x, y) = \dfrac{1 - x^2 - y^2}{\sqrt{|1 - x^2 - y^2|}}$

Rotatable Graph

Section 13.2 Limits and Continuity

- **Understand the definition of a neighborhood in the plane.**
- **Understand and use the definition of the limit of a function of two variables.**
- **Extend the concept of continuity to a function of two variables.**
- **Extend the concept of continuity to a function of three variables.**

SONYA KOVALEVSKY (1850–1891)

Much of the terminology used to define limits and continuity of a function of two or three variables was introduced by the German mathematician Karl Weierstrass (1815–1897). Weierstrass's rigorous approach to limits and other topics in calculus gained him the reputation as the "father of modern analysis." Weierstrass was a gifted teacher. One of his best-known students was the Russian mathematician Sonya Kovalevsky, who applied many of Weierstrass's techniques to problems in mathematical physics and became one of the first women to gain acceptance as a research mathematician.

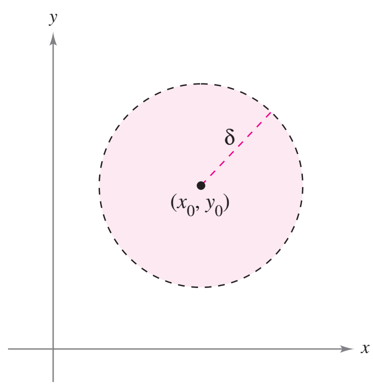
Neighborhoods in the Plane

In this section, you will study limits and continuity involving functions of two or three variables. The section begins with functions of two variables. At the end of the section, the concepts are extended to functions of three variables.

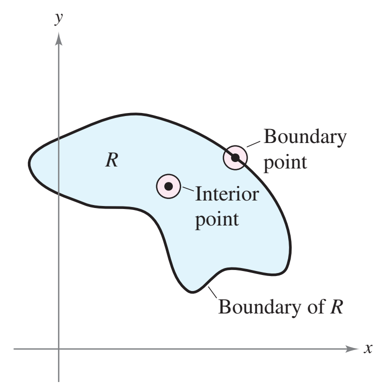
We begin our discussion of the limit of a function of two variables by defining a two-dimensional analog to an interval on the real line. Using the formula for the distance between two points (x, y) and (x_0, y_0) in the plane, you can define the **δ-neighborhood** about (x_0, y_0) to be the **disk** centered at (x_0, y_0) with radius $\delta > 0$

$$\left\{(x, y): \sqrt{(x - x_0)^2 + (y - y_0)^2} < \delta\right\} \qquad \text{Open disk}$$

as shown in Figure 13.18. When this formula contains the *less than* inequality, $<$, the disk is called **open,** and when it contains the *less than or equal to* inequality, $\leq$, the disk is called **closed.** This corresponds to the use of $<$ and $\leq$ to define open and closed intervals.



An open disk
Figure 13.18



The boundary and interior points of a region R
Figure 13.19

A point (x_0, y_0) in a plane region R is an **interior point** of R if there exists a δ-neighborhood about (x_0, y_0) that lies entirely in R, as shown in Figure 13.19. If every point in R is an interior point, then R is an **open region.** A point (x_0, y_0) is a **boundary point** of R if every open disk centered at (x_0, y_0) contains points inside R *and* points outside R. By definition, a region must contain its interior points, but it need not contain its boundary points. If a region contains all its boundary points, the region is **closed.** A region that contains some but not all of its boundary points is neither open nor closed.

FOR FURTHER INFORMATION For more information on Sonya Kovalevsky, see the article "S. Kovalevsky: A Mathematical Lesson" by Karen D. Rappaport in *The American Mathematical Monthly*.

MathArticle

Limit of a Function of Two Variables

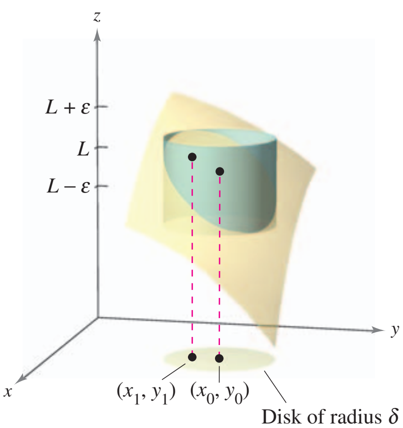
Definition of the Limit of a Function of Two Variables

Let f be a function of two variables defined, except possibly at (x_0, y_0), on an open disk centered at (x_0, y_0), and let L be a real number. Then

$$\lim_{(x, y) \to (x_0, y_0)} f(x, y) = L$$

if for each $\varepsilon > 0$ there corresponds a $\delta > 0$ such that

$$|f(x, y) - L| < \varepsilon \quad \text{whenever} \quad 0 < \sqrt{(x - x_0)^2 + (y - y_0)^2} < \delta.$$



For any (x, y) in the circle of radius δ, the value $f(x, y)$ lies between $L + \varepsilon$ and $L - \varepsilon$.
Figure 13.20

Rotatable Graph

NOTE Graphically, this definition of a limit implies that for any point $(x, y) \neq (x_0, y_0)$ in the disk of radius δ, the value $f(x, y)$ lies between $L + \varepsilon$ and $L - \varepsilon$, as shown in Figure 13.20.

The definition of the limit of a function of two variables is similar to the definition of the limit of a function of a single variable, yet there is a critical difference. To determine whether a function of a single variable has a limit, you need only test the approach from two directions—from the right and from the left. If the function approaches the same limit from the right and from the left, you can conclude that the limit exists. However, for a function of two variables, the statement

$$(x, y) \to (x_0, y_0)$$

means that the point (x, y) is allowed to approach (x_0, y_0) from any direction. If the value of

$$\lim_{(x, y) \to (x_0, y_0)} f(x, y)$$

is not the same for all possible approaches, or **paths,** to (x_0, y_0), the limit does not exist.

EXAMPLE 1 Verifying a Limit by the Definition

Show that

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

Solution Let $f(x, y) = x$ and $L = a$. You need to show that for each $\varepsilon > 0$, there exists a δ-neighborhood about (a, b) such that

$$|f(x, y) - L| = |x - a| < \varepsilon$$

whenever $(x, y) \neq (a, b)$ lies in the neighborhood. You can first observe that from

$$0 < \sqrt{(x - a)^2 + (y - b)^2} < \delta$$

it follows that

$$\begin{aligned} |f(x, y) - a| &= |x - a| \\ &= \sqrt{(x - a)^2} \\ &\leq \sqrt{(x - a)^2 + (y - b)^2} \\ &< \delta. \end{aligned}$$

So, you can choose $\delta = \varepsilon$, and the limit is verified.

Try It Exploration A

Limits of functions of several variables have the same properties regarding sums, differences, products, and quotients as do limits of functions of single variables. (See Theorem 1.2 in Section 1.3.) Some of these properties are used in the next example.

EXAMPLE 2 Verifying a Limit

Evaluate $\displaystyle\lim_{(x,y)\to(1,2)} \frac{5x^2y}{x^2+y^2}$.

Solution By using the properties of limits of products and sums, you obtain

$$\lim_{(x,y)\to(1,2)} 5x^2y = 5(1^2)(2) = 10$$

and

$$\lim_{(x,y)\to(1,2)} (x^2+y^2) = (1^2+2^2) = 5.$$

Because the limit of a quotient is equal to the quotient of the limits (and the denominator is not 0), you have

$$\lim_{(x,y)\to(1,2)} \frac{5x^2y}{x^2+y^2} = \frac{10}{5} = 2.$$

Try It Exploration A

EXAMPLE 3 Verifying a Limit

Evaluate $\displaystyle\lim_{(x,y)\to(0,0)} \frac{5x^2y}{x^2+y^2}$.

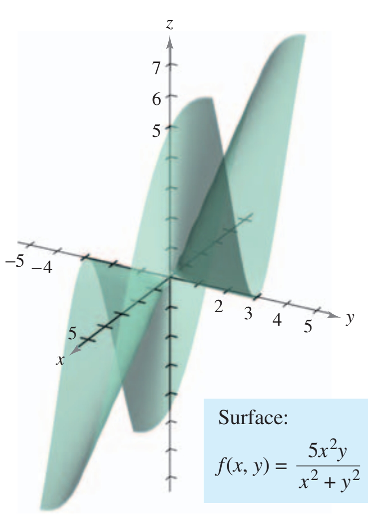


Figure 13.21

Rotatable Graph

Solution In this case, the limits of the numerator and of the denominator are both 0, and so you cannot determine the existence (or nonexistence) of a limit by taking the limits of the numerator and denominator separately and then dividing. However, from the graph of f in Figure 13.21, it seems reasonable that the limit might be 0. So, you can try applying the definition to $L = 0$. First, note that

$$|y| \le \sqrt{x^2+y^2} \qquad \text{and} \qquad \frac{x^2}{x^2+y^2} \le 1.$$

Then, in a δ-neighborhood about $(0, 0)$, you have $0 < \sqrt{x^2+y^2} < \delta$, and it follows that, for $(x, y) \ne (0, 0)$,

$$\begin{aligned}|f(x,y) - 0| &= \left|\frac{5x^2y}{x^2+y^2}\right| \\ &= 5|y|\left(\frac{x^2}{x^2+y^2}\right) \\ &\le 5|y| \\ &\le 5\sqrt{x^2+y^2} \\ &< 5\delta.\end{aligned}$$

So, you can choose $\delta = \varepsilon/5$ and conclude that

$$\lim_{(x,y)\to(0,0)} \frac{5x^2y}{x^2+y^2} = 0.$$

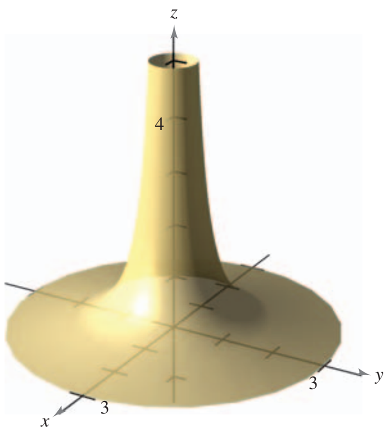
Try It Exploration A

For some functions, it is easy to recognize that a limit does not exist. For instance, it is clear that the limit

$$\lim_{(x,y)\to(0,0)} \frac{1}{x^2+y^2}$$

does not exist because the values of $f(x, y)$ increase without bound as (x, y) approaches $(0, 0)$ along *any* path (see Figure 13.22).

For other functions, it is not so easy to recognize that a limit does not exist. For instance, the next example describes a limit that does not exist because the function approaches different values along different paths.



$\displaystyle\lim_{(x,y)\to(0,0)} \frac{1}{x^2+y^2}$ does not exist.

Figure 13.22

Rotatable Graph

EXAMPLE 4 A Limit That Does Not Exist

Show that the following limit does not exist.

$$\lim_{(x,y)\to(0,0)} \left(\frac{x^2-y^2}{x^2+y^2}\right)^2$$

Solution The domain of the function given by

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

consists of all points in the xy-plane except for the point $(0, 0)$. To show that the limit as (x, y) approaches $(0, 0)$ does not exist, consider approaching $(0, 0)$ along two different "paths," as shown in Figure 13.23. Along the x-axis, every point is of the form $(x, 0)$, and the limit along this approach is

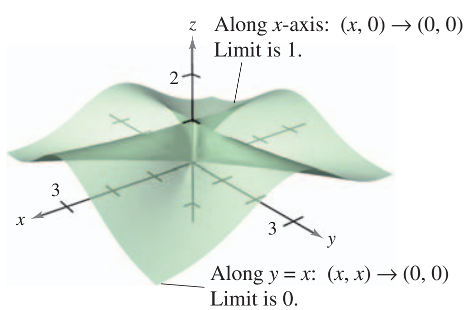
$$\lim_{(x,0)\to(0,0)} \left(\frac{x^2-0^2}{x^2+0^2}\right)^2 = \lim_{(x,0)\to(0,0)} 1^2 = 1. \qquad \text{Limit along } x\text{-axis}$$

However, if (x, y) approaches $(0, 0)$ along the line $y = x$, you obtain

$$\lim_{(x,x)\to(0,0)} \left(\frac{x^2-x^2}{x^2+x^2}\right)^2 = \lim_{(x,x)\to(0,0)} \left(\frac{0}{2x^2}\right)^2 = 0. \qquad \text{Limit along line } y = x$$

This means that in any open disk centered at $(0, 0)$ there are points (x, y) at which f takes on the value 1, and other points at which f takes on the value 0. For instance, $f(x, y) = 1$ at the points $(1, 0)$, $(0.1, 0)$, $(0.01, 0)$, and $(0.001, 0)$ and $f(x, y) = 0$ at the points $(1, 1)$, $(0.1, 0.1)$, $(0.01, 0.01)$, and $(0.001, 0.001)$. So, f does not have a limit as $(x, y) \to (0, 0)$.

NOTE In Example 4, you could conclude that the limit does not exist because you found two approaches that produced different limits. If two approaches had produced the same limit, you still could not have concluded that the limit exists. To form such a conclusion, you must show that the limit is the same along *all* possible approaches.



$\displaystyle\lim_{(x,y)\to(0,0)} \left(\frac{x^2-y^2}{x^2+y^2}\right)^2$ does not exist.

Figure 13.23

Rotatable Graph

Try It Exploration A Open Exploration

Continuity of a Function of Two Variables

Notice in Example 2 that the limit of $f(x, y) = 5x^2y/(x^2 + y^2)$ as $(x, y) \to (1, 2)$ can be evaluated by direct substitution. That is, the limit is $f(1, 2) = 2$. In such cases the function f is said to be **continuous** at the point $(1, 2)$.

NOTE This definition of continuity can be extended to *boundary points* of the open region R by considering a special type of limit in which (x, y) is allowed to approach (x_0, y_0) along paths lying in the region R. This notion is similar to that of one-sided limits, as discussed in Chapter 1.

Definition of Continuity of a Function of Two Variables

A function f of two variables is **continuous at a point (x_0, y_0)** in an open region R if $f(x_0, y_0)$ is equal to the limit of $f(x, y)$ as (x, y) approaches (x_0, y_0). That is,

$$\lim_{(x, y)\to(x_0, y_0)} f(x, y) = f(x_0, y_0).$$

The function f is **continuous in the open region R** if it is continuous at every point in R.

In Example 3, it was shown that the function

$$f(x, y) = \frac{5x^2y}{x^2 + y^2}$$

is not continuous at $(0, 0)$. However, because the limit at this point exists, you can remove the discontinuity by defining f at $(0, 0)$ as being equal to its limit there. Such a discontinuity is called **removable.** In Example 4, the function

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

was also shown not to be continuous at $(0, 0)$, but this discontinuity is **nonremovable.**

THEOREM 13.1 Continuous Functions of Two Variables

If k is a real number and f and g are continuous at (x_0, y_0), then the following functions are continuous at (x_0, y_0).

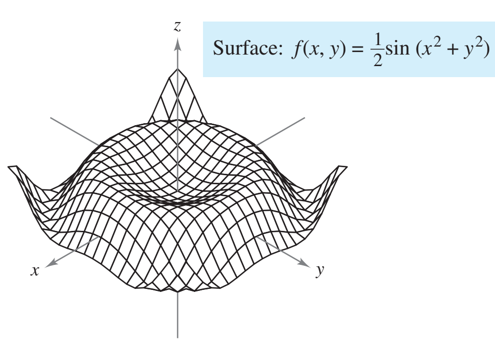
1. Scalar multiple: kf

2. Sum and difference: $f \pm g$

3. Product: fg

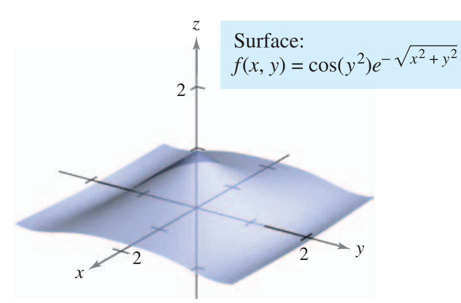
4. Quotient: f/g, if $g(x_0, y_0) \neq 0$

Theorem 13.1 establishes the continuity of *polynomial* and *rational* functions at every point in their domains. Furthermore, the continuity of other types of functions can be extended naturally from one to two variables. For instance, the functions whose graphs are shown in Figures 13.24 and 13.25 are continuous at every point in the plane.



The function f is continuous at every point in the plane.
Figure 13.24

Rotatable Graph



The function f is continuous at every point in the plane.
Figure 13.25

Rotatable Graph

EXPLORATION

Hold a spoon a foot or so from your eyes. Look at your image in the spoon. It should be upside down. Now, move the spoon closer and closer to one eye. At some point, your image will be right side up. Could it be that your image is being continuously deformed? Talk about this question and the general meaning of continuity with other members of your class. (This exploration was suggested by Irvin Roy Hentzel, Iowa State University.)

The next theorem states conditions under which a composite function is continuous.

THEOREM 13.2 Continuity of a Composite Function

If h is continuous at (x_0, y_0) and g is continuous at $h(x_0, y_0)$, then the composite function given by $(g \circ h)(x, y) = g(h(x, y))$ is continuous at (x_0, y_0). That is,

$$\lim_{(x, y)\to(x_0, y_0)} g(h(x, y)) = g(h(x_0, y_0)).$$

NOTE Note in Theorem 13.2 that h is a function of two variables and g is a function of one variable.

EXAMPLE 5 Testing for Continuity

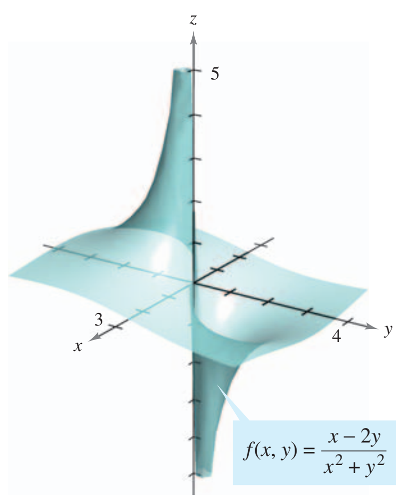
Discuss the continuity of each function.

a. $f(x, y) = \dfrac{x - 2y}{x^2 + y^2}$ **b.** $g(x, y) = \dfrac{2}{y - x^2}$

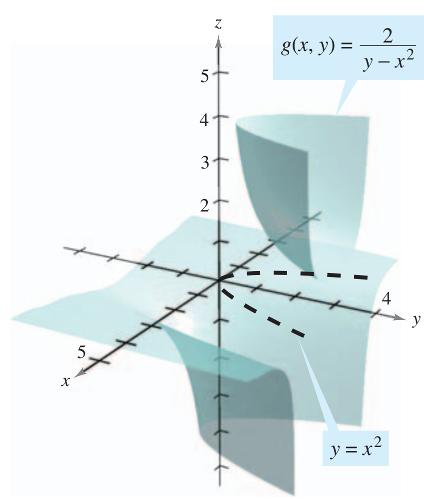
Solution

a. Because a rational function is continuous at every point in its domain, you can conclude that f is continuous at each point in the xy-plane except at $(0, 0)$, as shown in Figure 13.26.

b. The function given by $g(x, y) = 2/(y - x^2)$ is continuous except at the points at which the denominator is 0, $y - x^2 = 0$. So, you can conclude that the function is continuous at all points except those lying on the parabola $y = x^2$. Inside this parabola, you have $y > x^2$, and the surface represented by the function lies above the xy-plane, as shown in Figure 13.27. Outside the parabola, $y < x^2$, and the surface lies below the xy-plane.



The function f is not continuous at $(0, 0)$.
Figure 13.26



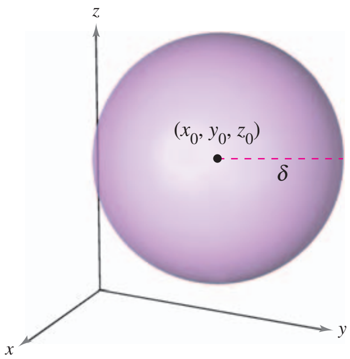
The function g is not continuous on the parabola $y = x^2$.
Figure 13.27

Rotatable Graph

Rotatable Graph

Try It

Exploration A



Open sphere in space
Figure 13.28

Rotatable Graph

Continuity of a Function of Three Variables

The preceding definitions of limits and continuity can be extended to functions of three variables by considering points (x, y, z) within the *open sphere*

$$(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2 < \delta^2. \quad \text{Open sphere}$$

The radius of this sphere is δ, and the sphere is centered at (x_0, y_0, z_0), as shown in Figure 13.28. A point (x_0, y_0, z_0) in a region R in space is an **interior point** of R if there exists a δ-sphere about (x_0, y_0, z_0) that lies entirely in R. If every point in R is an interior point, then R is called **open.**

Definition of Continuity of a Function of Three Variables

A function f of three variables is **continuous at a point** $(\boldsymbol{x_0, y_0, z_0})$ in an open region R if $f(x_0, y_0, z_0)$ is defined and is equal to the limit of $f(x, y, z)$ as (x, y, z) approaches (x_0, y_0, z_0). That is,

$$\lim_{(x, y, z) \to (x_0, y_0, z_0)} f(x, y, z) = f(x_0, y_0, z_0).$$

The function f is **continuous in the open region *R*** if it is continuous at every point in R.

EXAMPLE 6 Testing Continuity of a Function of Three Variables

The function

$$f(x, y, z) = \frac{1}{x^2 + y^2 - z}$$

is continuous at each point in space except at the points on the paraboloid given by $z = x^2 + y^2$.

Try It   Exploration A   Technology

Section 13.3 Partial Derivatives

- **Find and use partial derivatives of a function of two variables.**
- **Find and use partial derivatives of a function of three or more variables.**
- **Find higher-order partial derivatives of a function of two or three variables.**

JEAN LE ROND D'ALEMBERT (1717–1783)

The introduction of partial derivatives followed Newton's and Leibniz's work in calculus by several years. Between 1730 and 1760, Leonhard Euler and Jean Le Rond d'Alembert separately published several papers on dynamics, in which they established much of the theory of partial derivatives. These papers used functions of two or more variables to study problems involving equilibrium, fluid motion, and vibrating strings.

MathBio

Partial Derivatives of a Function of Two Variables

In applications of functions of several variables, the question often arises, "How will the value of a function be affected by a change in one of its independent variables?" You can answer this by considering the independent variables one at a time. For example, to determine the effect of a catalyst in an experiment, a chemist could conduct the experiment several times using varying amounts of the catalyst, while keeping constant other variables such as temperature and pressure. You can use a similar procedure to determine the rate of change of a function f with respect to one of its several independent variables. This process is called **partial differentiation,** and the result is referred to as the **partial derivative** of f with respect to the chosen independent variable.

Definition of Partial Derivatives of a Function of Two Variables

If $z = f(x, y)$, then the **first partial derivatives** of f with respect to x and y are the functions f_x and f_y defined by

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

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

provided the limits exist.

This definition indicates that if $z = f(x, y)$, then to find f_x you *consider y constant* and differentiate with respect to x. Similarly, to find f_y, you *consider x constant* and differentiate with respect to y.

EXAMPLE 1 Finding Partial Derivatives

Find the partial derivatives f_x and f_y for the function

$$f(x, y) = 3x - x^2y^2 + 2x^3y. \qquad \text{Original function}$$

Solution Considering y to be constant and differentiating with respect to x produces

$$f(x, y) = 3x - x^2y^2 + 2x^3y \qquad \text{Write original function.}$$

$$f_x(x, y) = 3 - 2xy^2 + 6x^2y. \qquad \text{Partial derivative with respect to } x$$

Considering x to be constant and differentiating with respect to y produces

$$f(x, y) = 3x - x^2y^2 + 2x^3y \qquad \text{Write original function.}$$

$$f_y(x, y) = -2x^2y + 2x^3. \qquad \text{Partial derivative with respect to } y$$

Try It Exploration A

Notation for First Partial Derivatives

For $z = f(x, y)$, the partial derivatives f_x and f_y are denoted by

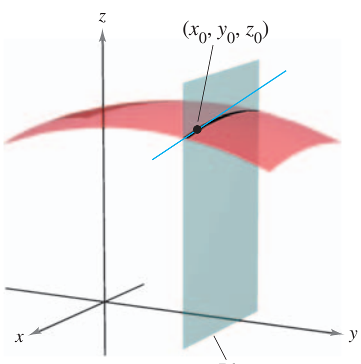
$$\frac{\partial}{\partial x} f(x, y) = f_x(x, y) = z_x = \frac{\partial z}{\partial x}$$

and

$$\frac{\partial}{\partial y} f(x, y) = f_y(x, y) = z_y = \frac{\partial z}{\partial y}.$$

The first partials evaluated at the point (a, b) are denoted by

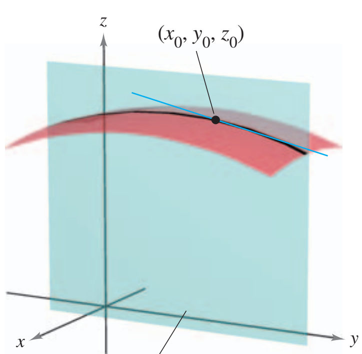
$$\left.\frac{\partial z}{\partial x}\right|_{(a,\, b)} = f_x(a, b) \qquad \text{and} \qquad \left.\frac{\partial z}{\partial y}\right|_{(a,\, b)} = f_y(a, b).$$



$\frac{\partial f}{\partial x}$ = slope in x-direction

Figure 13.29

Rotatable Graph



$\frac{\partial f}{\partial y}$ = slope in y-direction

Figure 13.30

Rotatable Graph

EXAMPLE 2 Finding and Evaluating Partial Derivatives

For $f(x, y) = xe^{x^2 y}$, find f_x and f_y, and evaluate each at the point $(1, \ln 2)$.

Solution Because

$$f_x(x, y) = xe^{x^2 y}(2xy) + e^{x^2 y} \qquad \text{Partial derivative with respect to } x$$

the partial derivative of f with respect to x at $(1, \ln 2)$ is

$$\begin{aligned} f_x(1, \ln 2) &= e^{\ln 2}(2 \ln 2) + e^{\ln 2} \\ &= 4 \ln 2 + 2. \end{aligned}$$

Because

$$\begin{aligned} f_y(x, y) &= xe^{x^2 y}(x^2) \\ &= x^3 e^{x^2 y} \qquad \text{Partial derivative with respect to } y \end{aligned}$$

the partial derivative of f with respect to y at $(1, \ln 2)$ is

$$\begin{aligned} f_y(1, \ln 2) &= e^{\ln 2} \\ &= 2. \end{aligned}$$

Try It | Exploration A | Technology

The partial derivatives of a function of two variables, $z = f(x, y)$, have a useful geometric interpretation. If $y = y_0$, then $z = f(x, y_0)$ represents the curve formed by intersecting the surface $z = f(x, y)$ with the plane $y = y_0$, as shown in Figure 13.29. Therefore,

$$f_x(x_0, y_0) = \lim_{\Delta x \to 0} \frac{f(x_0 + \Delta x, y_0) - f(x_0, y_0)}{\Delta x}$$

represents the slope of this curve at the point $(x_0, y_0, f(x_0, y_0))$. Note that both the curve and the tangent line lie in the plane $y = y_0$. Similarly,

$$f_y(x_0, y_0) = \lim_{\Delta y \to 0} \frac{f(x_0, y_0 + \Delta y) - f(x_0, y_0)}{\Delta y}$$

represents the slope of the curve given by the intersection of $z = f(x, y)$ and the plane $x = x_0$ at $(x_0, y_0, f(x_0, y_0))$, as shown in Figure 13.30.

Informally, the values of $\partial f/\partial x$ and $\partial f/\partial y$ at the point (x_0, y_0, z_0) denote the **slopes of the surface in the x- and y-directions,** respectively.

EXAMPLE 3 Finding the Slopes of a Surface in the *x*- and *y*-Directions

Find the slopes in the x-direction and in the y-direction of the surface given by

$$f(x, y) = -\frac{x^2}{2} - y^2 + \frac{25}{8}$$

at the point $\left(\frac{1}{2}, 1, 2\right)$.

Solution The partial derivatives of f with respect to x and y are

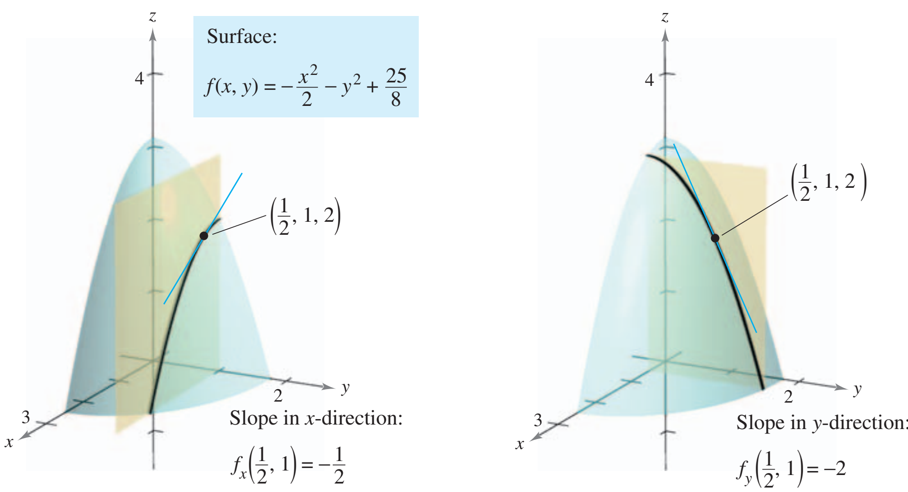
$$f_x(x, y) = -x \qquad \text{and} \qquad f_y(x, y) = -2y. \qquad \text{Partial derivatives}$$

So, in the x-direction, the slope is

$$f_x\left(\frac{1}{2}, 1\right) = -\frac{1}{2} \qquad \text{Figure 13.31(a)}$$

and in the y-direction, the slope is

$$f_y\left(\frac{1}{2}, 1\right) = -2. \qquad \text{Figure 13.31(b)}$$



(a) Rotatable Graph **(b)** Rotatable Graph

Figure 13.31

Try It | Exploration A | Open Exploration

EXAMPLE 4 Finding the Slopes of a Surface in the *x*- and *y*-Directions

Find the slopes of the surface given by

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

at the point $(1, 2, 1)$ in the x-direction and in the y-direction.

Solution The partial derivatives of f with respect to x and y are

$$f_x(x, y) = -2(x - 1) \qquad \text{and} \qquad f_y(x, y) = -2(y - 2). \qquad \text{Partial derivatives}$$

So, at the point $(1, 2, 1)$, the slopes in the x- and y-directions are

$$f_x(1, 2) = -2(1 - 1) = 0 \qquad \text{and} \qquad f_y(1, 2) = -2(2 - 2) = 0$$

as shown in Figure 13.32.

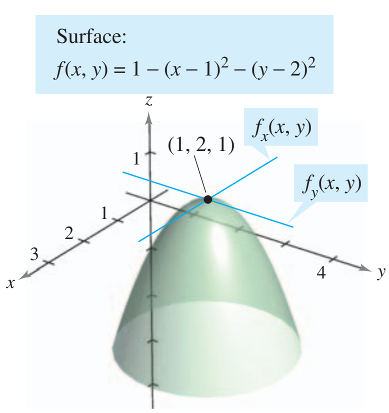


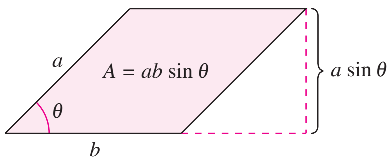
Figure 13.32

Rotatable Graph

Try It | Exploration A

No matter how many variables are involved, partial derivatives can be interpreted as *rates of change*.

EXAMPLE 5 Using Partial Derivatives to Find Rates of Change



The area of the parallelogram is $ab\sin\theta$.
Figure 13.33

The area of a parallelogram with adjacent sides a and b and included angle θ is given by $A = ab\sin\theta$, as shown in Figure 13.33.

a. Find the rate of change of A with respect to a for $a = 10$, $b = 20$, and $\theta = \dfrac{\pi}{6}$.

b. Find the rate of change of A with respect to θ for $a = 10$, $b = 20$, and $\theta = \dfrac{\pi}{6}$.

Solution

a. To find the rate of change of the area with respect to a, hold b and θ constant and differentiate with respect to a to obtain

$$\frac{\partial A}{\partial a} = b\sin\theta \qquad \text{Find partial with respect to } a.$$

$$\frac{\partial A}{\partial a} = 20\sin\frac{\pi}{6} = 10. \qquad \text{Substitute for } b \text{ and } \theta.$$

b. To find the rate of change of the area with respect to θ, hold a and b constant and differentiate with respect to θ to obtain

$$\frac{\partial A}{\partial \theta} = ab\cos\theta \qquad \text{Find partial with respect to } \theta.$$

$$\frac{\partial A}{\partial \theta} = 200\cos\frac{\pi}{6} = 100\sqrt{3}. \qquad \text{Substitute for } a,\, b, \text{ and } \theta.$$

Try It Exploration A

Partial Derivatives of a Function of Three or More Variables

The concept of a partial derivative can be extended naturally to functions of three or more variables. For instance, if $w = f(x, y, z)$, there are three partial derivatives, each of which is formed by holding two of the variables constant. That is, to define the partial derivative of w with respect to x, consider y and z to be constant and differentiate with respect to x. A similar process is used to find the derivatives of w with respect to y and with respect to z.

$$\frac{\partial w}{\partial x} = f_x(x, y, z) = \lim_{\Delta x \to 0} \frac{f(x + \Delta x, y, z) - f(x, y, z)}{\Delta x}$$

$$\frac{\partial w}{\partial y} = f_y(x, y, z) = \lim_{\Delta y \to 0} \frac{f(x, y + \Delta y, z) - f(x, y, z)}{\Delta y}$$

$$\frac{\partial w}{\partial z} = f_z(x, y, z) = \lim_{\Delta z \to 0} \frac{f(x, y, z + \Delta z) - f(x, y, z)}{\Delta z}$$

In general, if $w = f(x_1, x_2, \ldots, x_n)$, there are n partial derivatives denoted by

$$\frac{\partial w}{\partial x_k} = f_{x_k}(x_1, x_2, \ldots, x_n), \quad k = 1, 2, \ldots, n.$$

To find the partial derivative with respect to one of the variables, hold the other variables constant and differentiate with respect to the given variable.

EXAMPLE 6 Finding Partial Derivatives

a. To find the partial derivative of $f(x, y, z) = xy + yz^2 + xz$ with respect to z, consider x and y to be constant and obtain

$$\frac{\partial}{\partial z}[xy + yz^2 + xz] = 2yz + x.$$

b. To find the partial derivative of $f(x, y, z) = z\sin(xy^2 + 2z)$ with respect to z, consider x and y to be constant. Then, using the Product Rule, you obtain

$$\begin{aligned}\frac{\partial}{\partial z}[z\sin(xy^2 + 2z)] &= (z)\frac{\partial}{\partial z}[\sin(xy^2 + 2z)] + \sin(xy^2 + 2z)\frac{\partial}{\partial z}[z]\\ &= (z)[\cos(xy^2 + 2z)](2) + \sin(xy^2 + 2z)\\ &= 2z\cos(xy^2 + 2z) + \sin(xy^2 + 2z).\end{aligned}$$

c. To find the partial derivative of $f(x, y, z, w) = (x + y + z)/w$ with respect to w, consider x, y, and z to be constant and obtain

$$\frac{\partial}{\partial w}\left[\frac{x + y + z}{w}\right] = -\frac{x + y + z}{w^2}.$$

Try It   Exploration A

Higher-Order Partial Derivatives

As is true for ordinary derivatives, it is possible to take second, third, and higher partial derivatives of a function of several variables, provided such derivatives exist. Higher-order derivatives are denoted by the order in which the differentiation occurs. For instance, the function $z = f(x, y)$ has the following second partial derivatives.

1. Differentiate twice with respect to x:

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial x^2} = f_{xx}.$$

2. Differentiate twice with respect to y:

$$\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial y^2} = f_{yy}.$$

3. Differentiate first with respect to x and then with respect to y:

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

4. Differentiate first with respect to y and then with respect to x:

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial x\partial y} = f_{yx}.$$

The third and fourth cases are called **mixed partial derivatives.**

NOTE Note that the two types of notation for mixed partials have different conventions for indicating the order of differentiation.

$$\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial y\partial x} \quad \text{Right-to-left order}$$

$$(f_x)_y = f_{xy} \quad \text{Left-to-right order}$$

You can remember the order by observing that in both notations, you differentiate first with respect to the variable "nearest" f.

EXAMPLE 7 Finding Second Partial Derivatives

Find the second partial derivatives of $f(x, y) = 3xy^2 - 2y + 5x^2y^2$, and determine the value of $f_{xy}(-1, 2)$.

Solution Begin by finding the first partial derivatives with respect to x and y.

$$f_x(x, y) = 3y^2 + 10xy^2 \qquad \text{and} \qquad f_y(x, y) = 6xy - 2 + 10x^2y$$

Then, differentiate each of these with respect to x and y.

$$f_{xx}(x, y) = 10y^2 \qquad \text{and} \qquad f_{yy}(x, y) = 6x + 10x^2$$
$$f_{xy}(x, y) = 6y + 20xy \qquad \text{and} \qquad f_{yx}(x, y) = 6y + 20xy$$

At $(-1, 2)$, the value of f_{xy} is $f_{xy}(-1, 2) = 12 - 40 = -28$.

Try It   Exploration A

NOTE Notice in Example 7 that the two mixed partials are equal. Sufficient conditions for this occurrence are given in Theorem 13.3.

THEOREM 13.3 Equality of Mixed Partial Derivatives

If f is a function of x and y such that f_{xy} and f_{yx} are continuous on an open disk R, then, for every (x, y) in R,

$$f_{xy}(x, y) = f_{yx}(x, y).$$

Theorem 13.3 also applies to a function f of *three or more variables* so long as all second partial derivatives are continuous. For example, if $w = f(x, y, z)$ and all the second partial derivatives are continuous in an open region R, then at each point in R the order of differentiation in the mixed second partial derivatives is irrelevant. If the third partial derivatives of f are also continuous, the order of differentiation of the mixed third partial derivatives is irrelevant.

EXAMPLE 8 Finding Higher-Order Partial Derivatives

Show that $f_{xz} = f_{zx}$ and $f_{xzz} = f_{zxz} = f_{zzx}$ for the function given by

$$f(x, y, z) = ye^x + x \ln z.$$

Solution

First partials:

$$f_x(x, y, z) = ye^x + \ln z, \qquad f_z(x, y, z) = \frac{x}{z}$$

Second partials (note that the first two are equal):

$$f_{xz}(x, y, z) = \frac{1}{z}, \qquad f_{zx}(x, y, z) = \frac{1}{z}, \qquad f_{zz}(x, y, z) = -\frac{x}{z^2}$$

Third partials (note that all three are equal):

$$f_{xzz}(x, y, z) = -\frac{1}{z^2}, \qquad f_{zxz}(x, y, z) = -\frac{1}{z^2}, \qquad f_{zzx}(x, y, z) = -\frac{1}{z^2}$$

Try It   Exploration A   Exploration B

Section 13.4

Differentials

- **Understand the concepts of increments and differentials.**
- **Extend the concept of differentiability to a function of two variables.**
- **Use a differential as an approximation.**

Increments and Differentials

In this section, the concepts of increments and differentials are generalized to functions of two or more variables. Recall from Section 3.9 that for $y = f(x)$, the differential of y was defined as

$$dy = f'(x)\,dx.$$

Similar terminology is used for a function of two variables, $z = f(x, y)$. That is, Δx and Δy are the **increments of x and y,** and the **increment of z** is given by

$$\Delta z = f(x + \Delta x, y + \Delta y) - f(x, y).$$

Increment of z

Definition of Total Differential

If $z = f(x, y)$ and Δx and Δy are increments of x and y, then the **differentials** of the independent variables x and y are

$$dx = \Delta x \qquad \text{and} \qquad dy = \Delta y$$

and the **total differential** of the dependent variable z is

$$dz = \frac{\partial z}{\partial x}\,dx + \frac{\partial z}{\partial y}\,dy = f_x(x, y)\,dx + f_y(x, y)\,dy.$$

This definition can be extended to a function of three or more variables. For instance, if $w = f(x, y, z, u)$, then $dx = \Delta x$, $dy = \Delta y$, $dz = \Delta z$, $du = \Delta u$, and the total differential of w is

$$dw = \frac{\partial w}{\partial x}\,dx + \frac{\partial w}{\partial y}\,dy + \frac{\partial w}{\partial z}\,dz + \frac{\partial w}{\partial u}\,du.$$

EXAMPLE 1 Finding the Total Differential

Find the total differential for each function.

a. $z = 2x \sin y - 3x^2y^2$ **b.** $w = x^2 + y^2 + z^2$

Solution

a. The total differential dz for $z = 2x \sin y - 3x^2y^2$ is

$$\begin{aligned} dz &= \frac{\partial z}{\partial x}\,dx + \frac{\partial z}{\partial y}\,dy \\ &= (2 \sin y - 6xy^2)\,dx + (2x \cos y - 6x^2y)\,dy. \end{aligned}$$

Total differential dz

b. The total differential dw for $w = x^2 + y^2 + z^2$ is

$$\begin{aligned} dw &= \frac{\partial w}{\partial x}\,dx + \frac{\partial w}{\partial y}\,dy + \frac{\partial w}{\partial z}\,dz \\ &= 2x\,dx + 2y\,dy + 2z\,dz. \end{aligned}$$

Total differential dw

Try It Exploration A

Differentiability

In Section 3.9, you learned that for a *differentiable* function given by $y = f(x)$, you can use the differential $dy = f'(x)\,dx$ as an approximation (for small Δx) to the value $\Delta y = f(x + \Delta x) - f(x)$. When a similar approximation is possible for a function of two variables, the function is said to be **differentiable.** This is stated explicitly in the following definition.

Definition of Differentiability

A function f given by $z = f(x, y)$ is **differentiable** at (x_0, y_0) if Δz can be written in the form

$$\Delta z = f_x(x_0, y_0)\,\Delta x + f_y(x_0, y_0)\,\Delta y + \varepsilon_1 \Delta x + \varepsilon_2 \Delta y$$

where both ε_1 and $\varepsilon_2 \to 0$ as $(\Delta x, \Delta y) \to (0, 0)$. The function f is **differentiable in a region R** if it is differentiable at each point in R.

EXAMPLE 2 Showing That a Function Is Differentiable

Show that the function given by

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

is differentiable at every point in the plane.

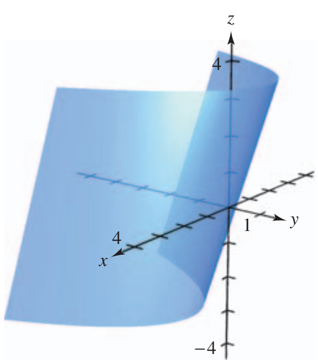


Figure 13.34

Rotatable Graph

Solution Letting $z = f(x, y)$, the increment of z at an arbitrary point (x, y) in the plane is

$$\begin{aligned}
\Delta z &= f(x + \Delta x, y + \Delta y) - f(x, y) && \text{Increment of } z\\
&= (x^2 + 2x\Delta x + \Delta x^2) + 3(y + \Delta y) - (x^2 + 3y)\\
&= 2x\Delta x + \Delta x^2 + 3\Delta y\\
&= 2x(\Delta x) + 3(\Delta y) + \Delta x(\Delta x) + 0(\Delta y)\\
&= f_x(x, y)\,\Delta x + f_y(x, y)\,\Delta y + \varepsilon_1 \Delta x + \varepsilon_2 \Delta y
\end{aligned}$$

where $\varepsilon_1 = \Delta x$ and $\varepsilon_2 = 0$. Because $\varepsilon_1 \to 0$ and $\varepsilon_2 \to 0$ as $(\Delta x, \Delta y) \to (0, 0)$, it follows that f is differentiable at every point in the plane. The graph of f is shown in Figure 13.34.

Try It   Exploration A

Be sure you see that the term "differentiable" is used differently for functions of two variables than for functions of one variable. A function of one variable is differentiable at a point if its derivative exists at the point. However, for a function of two variables, the existence of the partial derivatives f_x and f_y does not guarantee that the function is differentiable (see Example 5). The following theorem gives a *sufficient* condition for differentiability of a function of two variables. A proof of Theroem 13.4 is given in Appendix A.

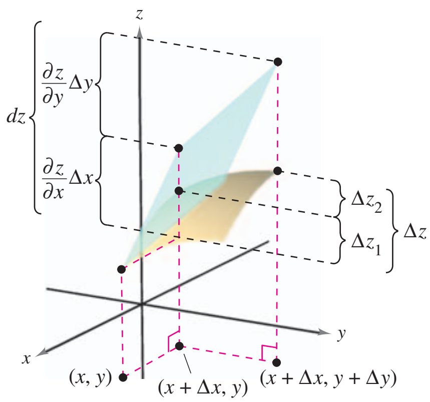
THEOREM 13.4 Sufficient Condition for Differentiability

If f is a function of x and y, where f_x and f_y are continuous in an open region R, then f is differentiable on R.

Approximation by Differentials

Theorem 13.4 tells you that you can choose $(x + \Delta x, y + \Delta y)$ close enough to (x, y) to make $\varepsilon_1 \Delta x$ and $\varepsilon_2 \Delta y$ insignificant. In other words, for small Δx and Δy, you can use the approximation

$$\Delta z \approx dz.$$



The exact change in z is Δz. This change can be approximated by the differential dz.
Figure 13.35

Rotatable Graph

This approximation is illustrated graphically in Figure 13.35. Recall that the partial derivatives $\partial z/\partial x$ and $\partial z/\partial y$ can be interpreted as the slopes of the surface in the x- and y-directions. This means that

$$dz = \frac{\partial z}{\partial x}\Delta x + \frac{\partial z}{\partial y}\Delta y$$

represents the change in height of a plane that is tangent to the surface at the point $(x, y, f(x, y))$. Because a plane in space is represented by a linear equation in the variables x, y, and z, the approximation of Δz by dz is called a **linear approximation.** You will learn more about this geometric interpretation in Section 13.7.

EXAMPLE 3 Using a Differential as an Approximation

Use the differential dz to approximate the change in $z = \sqrt{4 - x^2 - y^2}$ as (x, y) moves from the point $(1, 1)$ to the point $(1.01, 0.97)$. Compare this approximation with the exact change in z.

Solution Letting $(x, y) = (1, 1)$ and $(x + \Delta x, y + \Delta y) = (1.01, 0.97)$ produces $dx = \Delta x = 0.01$ and $dy = \Delta y = -0.03$. So, the change in z can be approximated by

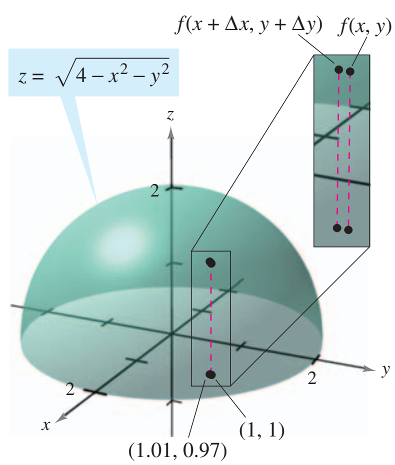
$$\Delta z \approx dz = \frac{\partial z}{\partial x}dx + \frac{\partial z}{\partial y}dy = \frac{-x}{\sqrt{4 - x^2 - y^2}}\Delta x + \frac{-y}{\sqrt{4 - x^2 - y^2}}\Delta y.$$

When $x = 1$ and $y = 1$, you have

$$\Delta z \approx -\frac{1}{\sqrt{2}}(0.01) - \frac{1}{\sqrt{2}}(-0.03) = \frac{0.02}{\sqrt{2}} = \sqrt{2}(0.01) \approx 0.0141.$$

In Figure 13.36 you can see that the exact change corresponds to the difference in the heights of two points on the surface of a hemisphere. This difference is given by

$$\begin{aligned}\Delta z &= f(1.01, 0.97) - f(1, 1)\\ &= \sqrt{4 - (1.01)^2 - (0.97)^2} - \sqrt{4 - 1^2 - 1^2} \approx 0.0137.\end{aligned}$$



As (x, y) moves from $(1, 1)$ to the point $(1.01, 0.97)$, the value of $f(x, y)$ changes by about 0.0137.
Figure 13.36

Rotatable Graph

Try It Exploration A Open Exploration

A function of three variables $w = f(x, y, z)$ is called **differentiable** at (x, y, z) provided that

$$\Delta w = f(x + \Delta x, y + \Delta y, z + \Delta z) - f(x, y, z)$$

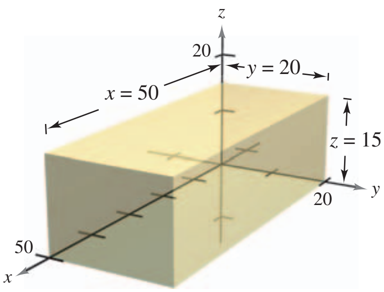
can be written in the form

$$\Delta w = f_x \Delta x + f_y \Delta y + f_z \Delta z + \varepsilon_1 \Delta x + \varepsilon_2 \Delta y + \varepsilon_3 \Delta z$$

where ε_1, ε_2, and $\varepsilon_3 \to 0$ as $(\Delta x, \Delta y, \Delta z) \to (0, 0, 0)$. With this definition of differentiability, Theorem 13.4 has the following extension for functions of three variables: If f is a function of x, y, and z, where f, f_x, f_y, and f_z are continuous in an open region R, then f is differentiable on R.

In Section 3.9, you used differentials to approximate the propagated error introduced by an error in measurement. This application of differentials is further illustrated in Example 4.

EXAMPLE 4 Error Analysis



Volume $= xyz$
Figure 13.37

Rotatable Graph

The possible error involved in measuring each dimension of a rectangular box is ± 0.1 millimeter. The dimensions of the box are $x = 50$ centimeters, $y = 20$ centimeters, and $z = 15$ centimeters, as shown in Figure 13.37. Use dV to estimate the propagated error and the relative error in the calculated volume of the box.

Solution The volume of the box is given by $V = xyz$, and so

$$\begin{aligned} dV &= \frac{\partial V}{\partial x}dx + \frac{\partial V}{\partial y}dy + \frac{\partial V}{\partial z}dz \\ &= yz\,dx + xz\,dy + xy\,dz. \end{aligned}$$

Using 0.1 millimeter $= 0.01$ centimeter, you have $dx = dy = dz = \pm 0.01$, and the propagated error is approximately

$$\begin{aligned} dV &= (20)(15)(\pm 0.01) + (50)(15)(\pm 0.01) + (50)(20)(\pm 0.01) \\ &= 300(\pm 0.01) + 750(\pm 0.01) + 1000(\pm 0.01) \\ &= 2050(\pm 0.01) = \pm 20.5 \text{ cubic centimeters.} \end{aligned}$$

Because the measured volume is

$$V = (50)(20)(15) = 15{,}000 \text{ cubic centimeters,}$$

the relative error, $\Delta V/V$, is approximately

$$\frac{\Delta V}{V} \approx \frac{dV}{V} = \frac{20.5}{15{,}000} \approx 0.14\%.$$

Try It Exploration A Exploration B

As is true for a function of a single variable, if a function in two or more variables is differentiable at a point, it is also continuous there.

THEOREM 13.5 Differentiability Implies Continuity

If a function of x and y is differentiable at (x_0, y_0), then it is continuous at (x_0, y_0).

Proof Let f be differentiable at (x_0, y_0), where $z = f(x, y)$. Then

$$\Delta z = [f_x(x_0, y_0) + \varepsilon_1]\,\Delta x + [f_y(x_0, y_0) + \varepsilon_2]\,\Delta y$$

where both ε_1 and $\varepsilon_2 \to 0$ as $(\Delta x, \Delta y) \to (0, 0)$. However, by definition, you know that Δz is given by

$$\Delta z = f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0).$$

Letting $x = x_0 + \Delta x$ and $y = y_0 + \Delta y$ produces

$$\begin{aligned} f(x, y) - f(x_0, y_0) &= [f_x(x_0, y_0) + \varepsilon_1]\,\Delta x + [f_y(x_0, y_0) + \varepsilon_2]\,\Delta y \\ &= [f_x(x_0, y_0) + \varepsilon_1](x - x_0) + [f_y(x_0, y_0) + \varepsilon_2](y - y_0). \end{aligned}$$

Taking the limit as $(x, y) \to (x_0, y_0)$, you have

$$\lim_{(x, y) \to (x_0, y_0)} f(x, y) = f(x_0, y_0)$$

which means that f is continuous at (x_0, y_0).

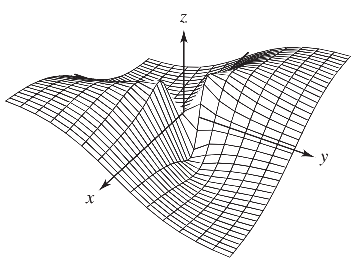
Remember that the existence of f_x and f_y is not sufficient to guarantee differentiability, as illustrated in the next example.

EXAMPLE 5 A Function That Is Not Differentiable

Show that $f_x(0, 0)$ and $f_y(0, 0)$ both exist, but that f is not differentiable at $(0, 0)$ where f is defined as

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

TECHNOLOGY Use a graphing utility to graph the function given in Example 5. For instance, the graph shown below was generated by *Mathematica*.



Generated by Mathematica

Solution You can show that f is not differentiable at $(0, 0)$ by showing that it is not continuous at this point. To see that f is not continuous at $(0, 0)$, look at the values of $f(x, y)$ along two different approaches to $(0, 0)$, as shown in Figure 13.38. Along the line $y = x$, the limit is

$$\lim_{(x, x)\to(0, 0)} f(x, y) = \lim_{(x, x)\to(0, 0)} \frac{-3x^2}{2x^2} = -\frac{3}{2}$$

whereas along $y = -x$ you have

$$\lim_{(x, -x)\to(0, 0)} f(x, y) = \lim_{(x, -x)\to(0, 0)} \frac{3x^2}{2x^2} = \frac{3}{2}.$$

So, the limit of $f(x, y)$ as $(x, y) \to (0, 0)$ does not exist, and you can conclude that f is not continuous at $(0, 0)$. Therefore, by Theorem 13.5, you know that f is not differentiable at $(0, 0)$. On the other hand, by the definition of the partial derivatives f_x and f_y, you have

$$f_x(0, 0) = \lim_{\Delta x\to 0} \frac{f(\Delta x, 0) - f(0, 0)}{\Delta x} = \lim_{\Delta x\to 0} \frac{0 - 0}{\Delta x} = 0$$

and

$$f_y(0, 0) = \lim_{\Delta y\to 0} \frac{f(0, \Delta y) - f(0, 0)}{\Delta y} = \lim_{\Delta y\to 0} \frac{0 - 0}{\Delta y} = 0.$$

So, the partial derivatives at $(0, 0)$ exist.

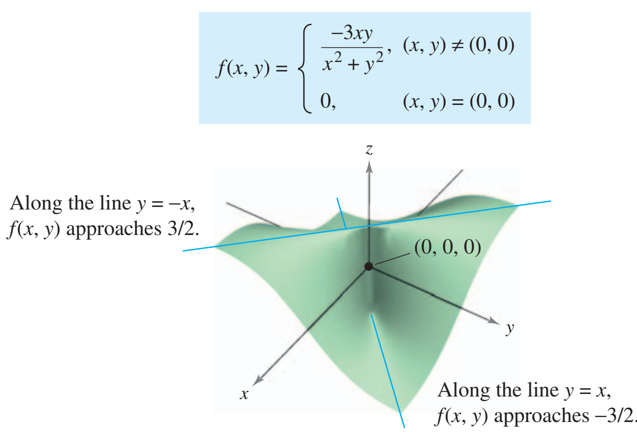


Figure 13.38

Rotatable Graph

Try It

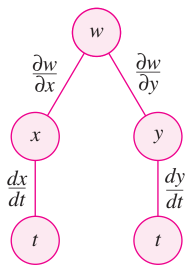
Exploration A

Section 13.5 Chain Rules for Functions of Several Variables

- **Use the Chain Rules for functions of several variables.**
- **Find partial derivatives implicitly.**

Chain Rules for Functions of Several Variables

Your work with differentials in the preceding section provides the basis for the extension of the Chain Rule to functions of two variables. There are two cases—the first case involves w as a function of x and y, where x and y are functions of a single independent variable t. (A proof of this theorem is given in Appendix A.)



Chain Rule: one independent variable w is a function of x and y, which are each functions of t. This diagram represents the derivative of w with respect to t.
Figure 13.39

THEOREM 13.6 Chain Rule: One Independent Variable

Let $w = f(x, y)$, where f is a differentiable function of x and y. If $x = g(t)$ and $y = h(t)$, where g and h are differentiable functions of t, then w is a differentiable function of t, and

$$\frac{dw}{dt} = \frac{\partial w}{\partial x}\frac{dx}{dt} + \frac{\partial w}{\partial y}\frac{dy}{dt}.$$

See Figure 13.39.

EXAMPLE 1 Using the Chain Rule with One Independent Variable

Let $w = x^2y - y^2$, where $x = \sin t$ and $y = e^t$. Find dw/dt when $t = 0$.

Solution By the Chain Rule for one independent variable, you have

$$\begin{aligned}\frac{dw}{dt} &= \frac{\partial w}{\partial x}\frac{dx}{dt} + \frac{\partial w}{\partial y}\frac{dy}{dt}\\ &= 2xy(\cos t) + (x^2 - 2y)e^t\\ &= 2(\sin t)(e^t)(\cos t) + (\sin^2 t - 2e^t)e^t\\ &= 2e^t \sin t \cos t + e^t \sin^2 t - 2e^{2t}.\end{aligned}$$

When $t = 0$, it follows that

$$\frac{dw}{dt} = -2.$$

Try It **Exploration A**

The Chain Rules presented in this section provide alternative techniques for solving many problems in single-variable calculus. For instance, in Example 1, you could have used single-variable techniques to find dw/dt by first writing w as a function of t,

$$\begin{aligned}w &= x^2y - y^2\\ &= (\sin t)^2(e^t) - (e^t)^2\\ &= e^t \sin^2 t - e^{2t}\end{aligned}$$

and then differentiating as usual.

$$\frac{dw}{dt} = 2e^t \sin t \cos t + e^t \sin^2 t - 2e^{2t}$$

The Chain Rule in Theorem 13.6 can be extended to any number of variables. For example, if each x_i is a differentiable function of a single variable t, then for

$$w = f(x_1, x_2, \ldots, x_n)$$

you have

$$\frac{dw}{dt} = \frac{\partial w}{\partial x_1}\frac{dx_1}{dt} + \frac{\partial w}{\partial x_2}\frac{dx_2}{dt} + \cdots + \frac{\partial w}{\partial x_n}\frac{dx_n}{dt}.$$

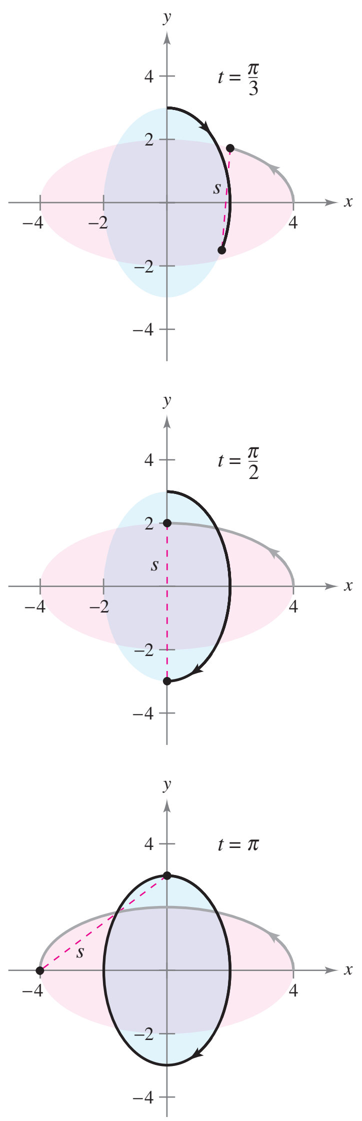
EXAMPLE 2 An Application of a Chain Rule to Related Rates

Two objects are traveling in elliptical paths given by the following parametric equations.

$x_1 = 4\cos t$ and $y_1 = 2\sin t$ First object

$x_2 = 2\sin 2t$ and $y_2 = 3\cos 2t$ Second object

At what rate is the distance between the two objects changing when $t = \pi$?



Paths of two objects traveling in elliptical orbits

Figure 13.40

Animation

Solution From Figure 13.40, you can see that the distance s between the two objects is given by

$$s = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

and that when $t = \pi$, you have $x_1 = -4$, $y_1 = 0$, $x_2 = 0$, $y_2 = 3$, and

$$s = \sqrt{(0 + 4)^2 + (3 - 0)^2} = 5.$$

When $t = \pi$, the partial derivatives of s are as follows.

$$\frac{\partial s}{\partial x_1} = \frac{-(x_2 - x_1)}{\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}} = -\frac{1}{5}(0 + 4) = -\frac{4}{5}$$

$$\frac{\partial s}{\partial y_1} = \frac{-(y_2 - y_1)}{\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}} = -\frac{1}{5}(3 - 0) = -\frac{3}{5}$$

$$\frac{\partial s}{\partial x_2} = \frac{(x_2 - x_1)}{\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}} = \frac{1}{5}(0 + 4) = \frac{4}{5}$$

$$\frac{\partial s}{\partial y_2} = \frac{(y_2 - y_1)}{\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}} = \frac{1}{5}(3 - 0) = \frac{3}{5}$$

When $t = \pi$, the derivatives of x_1, y_1, x_2, and y_2 are

$$\frac{dx_1}{dt} = -4\sin t = 0 \qquad \frac{dy_1}{dt} = 2\cos t = -2$$

$$\frac{dx_2}{dt} = 4\cos 2t = 4 \qquad \frac{dy_2}{dt} = -6\sin 2t = 0.$$

So, using the appropriate Chain Rule, you know that the distance is changing at a rate of

$$\begin{aligned}\frac{ds}{dt} &= \frac{\partial s}{\partial x_1}\frac{dx_1}{dt} + \frac{\partial s}{\partial y_1}\frac{dy_1}{dt} + \frac{\partial s}{\partial x_2}\frac{dx_2}{dt} + \frac{\partial s}{\partial y_2}\frac{dy_2}{dt}\\ &= \left(-\frac{4}{5}\right)(0) + \left(-\frac{3}{5}\right)(-2) + \left(\frac{4}{5}\right)(4) + \left(\frac{3}{5}\right)(0)\\ &= \frac{22}{5}.\end{aligned}$$

Try It

Exploration A

In Example 2, note that s is the function of four *intermediate* variables, x_1, y_1, x_2, and y_2, each of which is a function of a single variable t. Another type of composite function is one in which the intermediate variables are themselves functions of more than one variable. For instance, if $w = f(x, y)$, where $x = g(s, t)$ and $y = h(s, t)$, it follows that w is a function of s and t, and you can consider the partial derivatives of w with respect to s and t. One way to find these partial derivatives is to write w as a function of s and t explicitly by substituting the equations $x = g(s, t)$ and $y = h(s, t)$ into the equation $w = f(x, y)$. Then you can find the partial derivatives in the usual way, as demonstrated in the next example.

EXAMPLE 3 Finding Partial Derivatives by Substitution

Find $\partial w/\partial s$ and $\partial w/\partial t$ for $w = 2xy$, where $x = s^2 + t^2$ and $y = s/t$.

Solution Begin by substituting $x = s^2 + t^2$ and $y = s/t$ into the equation $w = 2xy$ to obtain

$$w = 2xy = 2(s^2 + t^2)\left(\frac{s}{t}\right) = 2\left(\frac{s^3}{t} + st\right).$$

Then, to find $\partial w/\partial s$, hold t constant and differentiate with respect to s.

$$\begin{aligned}\frac{\partial w}{\partial s} &= 2\left(\frac{3s^2}{t} + t\right)\\ &= \frac{6s^2 + 2t^2}{t}\end{aligned}$$

Similarly, to find $\partial w/\partial t$, hold s constant and differentiate with respect to t to obtain

$$\begin{aligned}\frac{\partial w}{\partial t} &= 2\left(-\frac{s^3}{t^2} + s\right)\\ &= 2\left(\frac{-s^3 + st^2}{t^2}\right)\\ &= \frac{2st^2 - 2s^3}{t^2}.\end{aligned}$$

Try It Exploration A

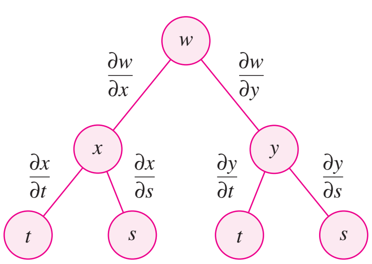
Theorem 13.7 gives an alternative method for finding the partial derivatives in Example 3, without explicitly writing w as a function of s and t.

THEOREM 13.7 Chain Rule: Two Independent Variables

Let $w = f(x, y)$, where f is a differentiable function of x and y. If $x = g(s, t)$ and $y = h(s, t)$ such that the first partials $\partial x/\partial s$, $\partial x/\partial t$, $\partial y/\partial s$, and $\partial y/\partial t$ all exist, then $\partial w/\partial s$ and $\partial w/\partial t$ exist and are given by

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

Proof To obtain $\partial w/\partial s$, hold t constant and apply Theorem 13.6 to obtain the desired result. Similarly, for $\partial w/\partial t$ hold s constant and apply Theorem 13.6.



Chain Rule: two independent variables
Figure 13.41

NOTE The Chain Rule in this theorem is shown schematically in Figure 13.41.

EXAMPLE 4 The Chain Rule with Two Independent Variables

Use the Chain Rule to find $\partial w/\partial s$ and $\partial w/\partial t$ for

$$w = 2xy$$

where $x = s^2 + t^2$ and $y = s/t$.

Solution Note that these same partials were found in Example 3. This time, using Theorem 13.7, you can hold t constant and differentiate with respect to s to obtain

$$\begin{aligned}
\frac{\partial w}{\partial s} &= \frac{\partial w}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial s} \\
&= 2y(2s) + 2x\left(\frac{1}{t}\right) \\
&= 2\left(\frac{s}{t}\right)(2s) + 2(s^2 + t^2)\left(\frac{1}{t}\right) && \text{Substitute } (s/t) \text{ for } y \text{ and } s^2 + t^2 \text{ for } x. \\
&= \frac{4s^2}{t} + \frac{2s^2 + 2t^2}{t} \\
&= \frac{6s^2 + 2t^2}{t}.
\end{aligned}$$

Similarly, holding s constant gives

$$\begin{aligned}
\frac{\partial w}{\partial t} &= \frac{\partial w}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial t} \\
&= 2y(2t) + 2x\left(\frac{-s}{t^2}\right) \\
&= 2\left(\frac{s}{t}\right)(2t) + 2(s^2 + t^2)\left(\frac{-s}{t^2}\right) && \text{Substitute } (s/t) \text{ for } y \text{ and } s^2 + t^2 \text{ for } x. \\
&= 4s - \frac{2s^3 + 2st^2}{t^2} \\
&= \frac{4st^2 - 2s^3 - 2st^2}{t^2} \\
&= \frac{2st^2 - 2s^3}{t^2}.
\end{aligned}$$

Try It | Open Exploration

The Chain Rule in Theorem 13.7 can also be extended to any number of variables. For example, if w is a differentiable function of the n variables $x_1, x_2, \ldots, x_n$, where each x_i is a differentiable function of the m variables $t_1, t_2, \ldots, t_m$, then for

$$w = f(x_1, x_2, \ldots, x_n)$$

you obtain the following.

$$\begin{aligned}
\frac{\partial w}{\partial t_1} &= \frac{\partial w}{\partial x_1}\frac{\partial x_1}{\partial t_1} + \frac{\partial w}{\partial x_2}\frac{\partial x_2}{\partial t_1} + \cdots + \frac{\partial w}{\partial x_n}\frac{\partial x_n}{\partial t_1} \\
\frac{\partial w}{\partial t_2} &= \frac{\partial w}{\partial x_1}\frac{\partial x_1}{\partial t_2} + \frac{\partial w}{\partial x_2}\frac{\partial x_2}{\partial t_2} + \cdots + \frac{\partial w}{\partial x_n}\frac{\partial x_n}{\partial t_2} \\
&\vdots \\
\frac{\partial w}{\partial t_m} &= \frac{\partial w}{\partial x_1}\frac{\partial x_1}{\partial t_m} + \frac{\partial w}{\partial x_2}\frac{\partial x_2}{\partial t_m} + \cdots + \frac{\partial w}{\partial x_n}\frac{\partial x_n}{\partial t_m}
\end{aligned}$$

EXAMPLE 5 The Chain Rule for a Function of Three Variables

Find $\partial w/\partial s$ and $\partial w/\partial t$ when $s = 1$ and $t = 2\pi$ for the function given by

$$w = xy + yz + xz$$

where $x = s\cos t$, $y = s\sin t$, and $z = t$.

Solution By extending the result of Theorem 13.7, you have

$$\begin{aligned}\frac{\partial w}{\partial s} &= \frac{\partial w}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial s} + \frac{\partial w}{\partial z}\frac{\partial z}{\partial s}\\ &= (y+z)(\cos t) + (x+z)(\sin t) + (y+x)(0)\\ &= (y+z)(\cos t) + (x+z)(\sin t).\end{aligned}$$

When $s = 1$ and $t = 2\pi$, you have $x = 1$, $y = 0$, and $z = 2\pi$. So, $\partial w/\partial s = (0 + 2\pi)(1) + (1 + 2\pi)(0) = 2\pi$. Furthermore,

$$\begin{aligned}\frac{\partial w}{\partial t} &= \frac{\partial w}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial t} + \frac{\partial w}{\partial z}\frac{\partial z}{\partial t}\\ &= (y+z)(-s\sin t) + (x+z)(s\cos t) + (y+x)(1)\end{aligned}$$

and for $s = 1$ and $t = 2\pi$, it follows that

$$\begin{aligned}\frac{\partial w}{\partial t} &= (0+2\pi)(0) + (1+2\pi)(1) + (0+1)(1)\\ &= 2 + 2\pi.\end{aligned}$$

Try It Exploration A

Implicit Partial Differentiation

This section concludes with an application of the Chain Rule to determine the derivative of a function defined *implicitly*. Suppose that x and y are related by the equation $F(x, y) = 0$, where it is assumed that $y = f(x)$ is a differentiable function of x. To find dy/dx, you could use the techniques discussed in Section 2.5. However, you will see that the Chain Rule provides a convenient alternative. If you consider the function given by

$$w = F(x, y) = F(x, f(x))$$

you can apply Theorem 13.6 to obtain

$$\frac{dw}{dx} = F_x(x, y)\frac{dx}{dx} + F_y(x, y)\frac{dy}{dx}.$$

Because $w = F(x, y) = 0$ for all x in the domain of f, you know that $dw/dx = 0$ and you have

$$F_x(x, y)\frac{dx}{dx} + F_y(x, y)\frac{dy}{dx} = 0.$$

Now, if $F_y(x, y) \neq 0$, you can use the fact that $dx/dx = 1$ to conclude that

$$\frac{dy}{dx} = -\frac{F_x(x, y)}{F_y(x, y)}.$$

A similar procedure can be used to find the partial derivatives of functions of several variables that are defined implicitly.

THEOREM 13.8 Chain Rule: Implicit Differentiation

If the equation $F(x, y) = 0$ defines y implicitly as a differentiable function of x, then

$$\frac{dy}{dx} = -\frac{F_x(x, y)}{F_y(x, y)}, \qquad F_y(x, y) \neq 0.$$

If the equation $F(x, y, z) = 0$ defines z implicitly as a differentiable function of x and y, then

$$\frac{\partial z}{\partial x} = -\frac{F_x(x, y, z)}{F_z(x, y, z)} \quad \text{and} \quad \frac{\partial z}{\partial y} = -\frac{F_y(x, y, z)}{F_z(x, y, z)}, \quad F_z(x, y, z) \neq 0.$$

This theorem can be extended to differentiable functions defined implicitly with any number of variables.

EXAMPLE 6 Finding a Derivative Implicitly

Find dy/dx, given $y^3 + y^2 - 5y - x^2 + 4 = 0$.

Solution Begin by defining a function F as

$$F(x, y) = y^3 + y^2 - 5y - x^2 + 4.$$

Then, using Theorem 13.8, you have

$$F_x(x, y) = -2x \quad \text{and} \quad F_y(x, y) = 3y^2 + 2y - 5$$

and it follows that

$$\frac{dy}{dx} = -\frac{F_x(x, y)}{F_y(x, y)} = \frac{-(-2x)}{3y^2 + 2y - 5} = \frac{2x}{3y^2 + 2y - 5}.$$

Try It Exploration A Exploration B

NOTE Compare the solution of Example 6 with the solution of Example 2 in Section 2.5.

EXAMPLE 7 Finding Partial Derivatives Implicitly

Find $\partial z/\partial x$ and $\partial z/\partial y$, given $3x^2z - x^2y^2 + 2z^3 + 3yz - 5 = 0$.

Solution To apply Theorem 13.8, let

$$F(x, y, z) = 3x^2z - x^2y^2 + 2z^3 + 3yz - 5.$$

Then

$$F_x(x, y, z) = 6xz - 2xy^2$$
$$F_y(x, y, z) = -2x^2y + 3z$$
$$F_z(x, y, z) = 3x^2 + 6z^2 + 3y$$

and you obtain

$$\frac{\partial z}{\partial x} = -\frac{F_x(x, y, z)}{F_z(x, y, z)} = \frac{2xy^2 - 6xz}{3x^2 + 6z^2 + 3y}$$

$$\frac{\partial z}{\partial y} = -\frac{F_y(x, y, z)}{F_z(x, y, z)} = \frac{2x^2y - 3z}{3x^2 + 6z^2 + 3y}.$$

Try It Exploration A

Section 13.6 Directional Derivatives and Gradients

- **Find and use directional derivatives of a function of two variables.**
- **Find the gradient of a function of two variables.**
- **Use the gradient of a function of two variables in applications.**
- **Find directional derivatives and gradients of functions of three variables.**

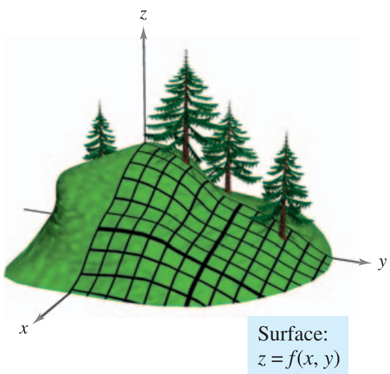


Figure 13.42

Rotatable Graph

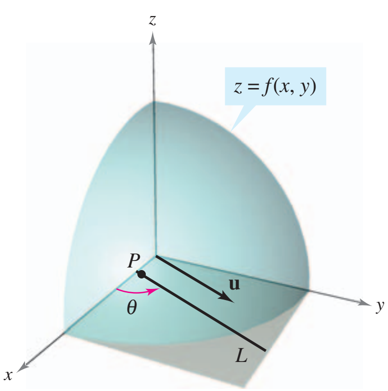


Figure 13.43

Rotatable Graph

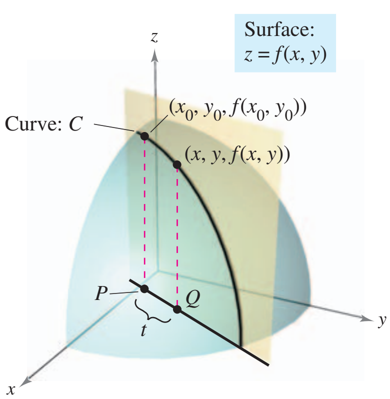


Figure 13.44

Rotatable Graph

Directional Derivative

You are standing on the hillside pictured in Figure 13.42 and want to determine the hill's incline toward the z-axis. If the hill were represented by $z = f(x, y)$, you would already know how to determine the slopes in two different directions—the slope in the y-direction would be given by the partial derivative $f_y(x, y)$, and the slope in the x-direction would be given by the partial derivative $f_x(x, y)$. In this section, you will see that these two partial derivatives can be used to find the slope in *any* direction.

To determine the slope at a point on a surface, you will define a new type of derivative called a **directional derivative.** Begin by letting $z = f(x, y)$ be a *surface* and $P(x_0, y_0)$ a *point* in the domain of f, as shown in Figure 13.43. The "direction" of the directional derivative is given by a unit vector

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

where θ is the angle the vector makes with the positive x-axis. To find the desired slope, reduce the problem to two dimensions by intersecting the surface with a vertical plane passing through the point P and parallel to $\mathbf{u}$, as shown in Figure 13.44. This vertical plane intersects the surface to form a curve C. The slope of the surface at $(x_0, y_0, f(x_0, y_0))$ in the direction of $\mathbf{u}$ is defined as the slope of the curve C at that point.

Informally, you can write the slope of the curve C as a limit that looks much like those used in single-variable calculus. The vertical plane used to form C intersects the xy-plane in a line L, represented by the parametric equations

$$x = x_0 + t\cos\theta$$

and

$$y = y_0 + t\sin\theta$$

so that for any value of t, the point $Q(x, y)$ lies on the line L. For each of the points P and Q, there is a corresponding point on the surface.

$(x_0, y_0, f(x_0, y_0))$ Point above P

$(x, y, f(x, y))$ Point above Q

Moreover, because the distance between P and Q is

$$\sqrt{(x - x_0)^2 + (y - y_0)^2} = \sqrt{(t\cos\theta)^2 + (t\sin\theta)^2} = |t|$$

you can write the slope of the secant line through $(x_0, y_0, f(x_0, y_0))$ and $(x, y, f(x, y))$ as

$$\frac{f(x, y) - f(x_0, y_0)}{t} = \frac{f(x_0 + t\cos\theta, y_0 + t\sin\theta) - f(x_0, y_0)}{t}.$$

Finally, by letting t approach 0, you arrive at the following definition.

Definition of Directional Derivative

Let f be a function of two variables x and y and let $\mathbf{u} = \cos\theta\,\mathbf{i} + \sin\theta\,\mathbf{j}$ be a unit vector. Then the **directional derivative of f in the direction of $\mathbf{u}$,** denoted by $D_{\mathbf{u}}f$, is

$$D_{\mathbf{u}}f(x, y) = \lim_{t\to 0} \frac{f(x + t\cos\theta, y + t\sin\theta) - f(x, y)}{t}$$

provided this limit exists.

Calculating directional derivatives by this definition is similar to finding the derivative of a function of one variable by the limit process (given in Section 2.1). A simpler "working" formula for finding directional derivatives involves the partial derivatives f_x and f_y.

THEOREM 13.9 Directional Derivative

If f is a differentiable function of x and y, then the directional derivative of f in the direction of the unit vector $\mathbf{u} = \cos\theta\,\mathbf{i} + \sin\theta\,\mathbf{j}$ is

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

Proof For a fixed point (x_0, y_0), let $x = x_0 + t\cos\theta$ and let $y = y_0 + t\sin\theta$. Then, let $g(t) = f(x, y)$. Because f is differentiable, you can apply the Chain Rule given in Theorem 13.7 to obtain

$$g'(t) = f_x(x, y)x'(t) + f_y(x, y)y'(t) = f_x(x, y)\cos\theta + f_y(x, y)\sin\theta.$$

If $t = 0$, then $x = x_0$ and $y = y_0$, so

$$g'(0) = f_x(x_0, y_0)\cos\theta + f_y(x_0, y_0)\sin\theta.$$

By the definition of $g'(t)$, it is also true that

$$\begin{aligned} g'(0) &= \lim_{t\to 0} \frac{g(t) - g(0)}{t} \\ &= \lim_{t\to 0} \frac{f(x_0 + t\cos\theta, y_0 + t\sin\theta) - f(x_0, y_0)}{t}. \end{aligned}$$

Consequently, $D_{\mathbf{u}}f(x_0, y_0) = f_x(x_0, y_0)\cos\theta + f_y(x_0, y_0)\sin\theta$.

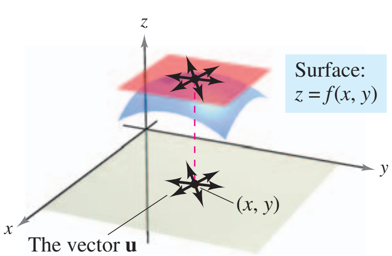


Figure 13.45

Rotatable Graph

There are infinitely many directional derivatives to a surface at a given point—one for each direction specified by $\mathbf{u}$, as shown in Figure 13.45. Two of these are the partial derivatives f_x and f_y.

1. Direction of positive x-axis ($\theta = 0$): $\mathbf{u} = \cos 0\,\mathbf{i} + \sin 0\,\mathbf{j} = \mathbf{i}$

$$D_{\mathbf{i}}f(x, y) = f_x(x, y)\cos 0 + f_y(x, y)\sin 0 = f_x(x, y)$$

2. Direction of positive y-axis ($\theta = \pi/2$): $\mathbf{u} = \cos\frac{\pi}{2}\,\mathbf{i} + \sin\frac{\pi}{2}\,\mathbf{j} = \mathbf{j}$

$$D_{\mathbf{j}}f(x, y) = f_x(x, y)\cos\frac{\pi}{2} + f_y(x, y)\sin\frac{\pi}{2} = f_y(x, y)$$

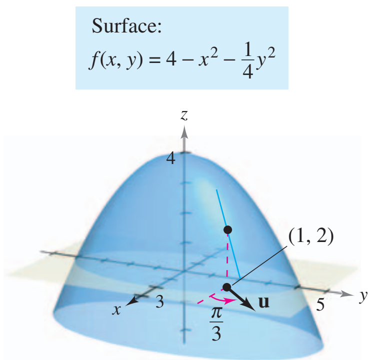


Figure 13.46

Rotatable Graph

NOTE Note in Figure 13.46 that you can interpret the directional derivative as giving the slope of the surface at the point (1, 2, 2) in the direction of the unit vector **u**.

EXAMPLE 1 Finding a Directional Derivative

Find the directional derivative of

$$f(x, y) = 4 - x^2 - \tfrac{1}{4}y^2 \qquad \text{Surface}$$

at (1, 2) in the direction of

$$\mathbf{u} = \left(\cos\frac{\pi}{3}\right)\mathbf{i} + \left(\sin\frac{\pi}{3}\right)\mathbf{j}. \qquad \text{Direction}$$

Solution Because f_x and f_y are continuous, f is differentiable, and you can apply Theorem 13.9.

$$\begin{aligned} D_{\mathbf{u}} f(x, y) &= f_x(x, y)\cos\theta + f_y(x, y)\sin\theta \\ &= (-2x)\cos\theta + \left(-\frac{y}{2}\right)\sin\theta \end{aligned}$$

Evaluating at $\theta = \pi/3$, $x = 1$, and $y = 2$ produces

$$\begin{aligned} D_{\mathbf{u}} f(1, 2) &= (-2)\left(\frac{1}{2}\right) + (-1)\left(\frac{\sqrt{3}}{2}\right) \\ &= -1 - \frac{\sqrt{3}}{2} \\ &\approx -1.866. \end{aligned} \qquad \text{See Figure 13.46.}$$

Try It Exploration A

You have been specifying direction by a unit vector **u**. If the direction is given by a vector whose length is not 1, you must normalize the vector before applying the formula in Theorem 13.9.

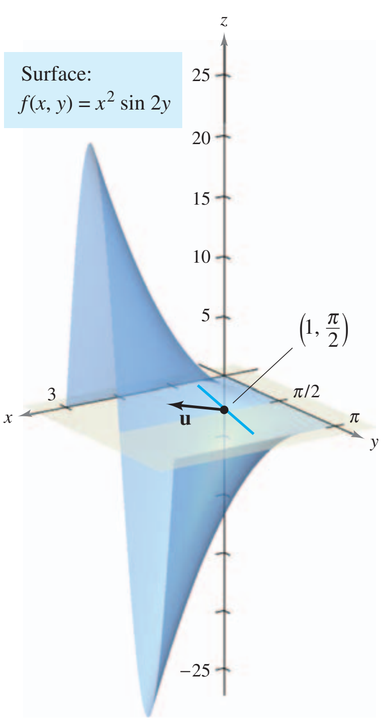


Figure 13.47

Rotatable Graph

EXAMPLE 2 Finding a Directional Derivative

Find the directional derivative of

$$f(x, y) = x^2 \sin 2y \qquad \text{Surface}$$

at $(1, \pi/2)$ in the direction of

$$\mathbf{v} = 3\mathbf{i} - 4\mathbf{j}. \qquad \text{Direction}$$

Solution Because f_x and f_y are continuous, f is differentiable, and you can apply Theorem 13.9. Begin by finding a unit vector in the direction of **v**.

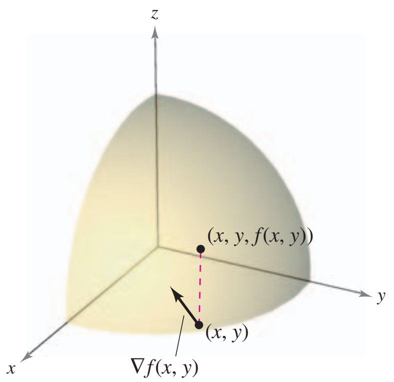
$$\mathbf{u} = \frac{\mathbf{v}}{\|\mathbf{v}\|} = \frac{3}{5}\mathbf{i} - \frac{4}{5}\mathbf{j} = \cos\theta\,\mathbf{i} + \sin\theta\,\mathbf{j}$$

Using this unit vector, you have

$$\begin{aligned} D_{\mathbf{u}} f(x, y) &= (2x\sin 2y)(\cos\theta) + (2x^2\cos 2y)(\sin\theta) \\ D_{\mathbf{u}} f\left(1, \frac{\pi}{2}\right) &= (2\sin\pi)\left(\frac{3}{5}\right) + (2\cos\pi)\left(-\frac{4}{5}\right) \\ &= (0)\left(\frac{3}{5}\right) + (-2)\left(-\frac{4}{5}\right) \\ &= \frac{8}{5}. \end{aligned} \qquad \text{See Figure 13.47.}$$

Try It Open Exploration

The Gradient of a Function of Two Variables



The gradient of f is a vector in the xy-plane
Figure 13.48

Rotatable Graph

The **gradient** of a function of two variables is a vector-valued function of two variables. This function has many important uses, some of which are described later in this section.

Definition of Gradient of a Function of Two Variables

Let $z = f(x, y)$ be a function of x and y such that f_x and f_y exist. Then the **gradient of f,** denoted by $\nabla f(x, y)$, is the vector

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

∇f is read as "del f." Another notation for the gradient is **grad** $f(x, y)$. In Figure 13.48, note that for each (x, y), the gradient $\nabla f(x, y)$ is a vector in the plane (not a vector in space).

NOTE No value is assigned to the symbol ∇ by itself. It is an operator in the same sense that d/dx is an operator. When ∇ operates on $f(x, y)$, it produces the vector $\nabla f(x, y)$.

EXAMPLE 3 Finding the Gradient of a Function

Find the gradient of $f(x, y) = y \ln x + xy^2$ at the point $(1, 2)$.

Solution Using

$$f_x(x, y) = \frac{y}{x} + y^2 \qquad \text{and} \qquad f_y(x, y) = \ln x + 2xy$$

you have

$$\nabla f(x, y) = \left(\frac{y}{x} + y^2\right)\mathbf{i} + (\ln x + 2xy)\mathbf{j}.$$

At the point $(1, 2)$, the gradient is

$$\begin{aligned}\nabla f(1, 2) &= \left(\frac{2}{1} + 2^2\right)\mathbf{i} + [\ln 1 + 2(1)(2)]\mathbf{j}\\ &= 6\mathbf{i} + 4\mathbf{j}.\end{aligned}$$

Try It   Exploration A

Because the gradient of f is a vector, you can write the directional derivative of f in the direction of $\mathbf{u}$ as

$$D_{\mathbf{u}} f(x, y) = [f_x(x, y)\mathbf{i} + f_y(x, y)\mathbf{j}] \cdot [\cos\theta\mathbf{i} + \sin\theta\mathbf{j}].$$

In other words, the directional derivative is the dot product of the gradient and the direction vector. This useful result is summarized in the following theorem.

THEOREM 13.10 Alternative Form of the Directional Derivative

If f is a differentiable function of x and y, then the directional derivative of f in the direction of the unit vector $\mathbf{u}$ is

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

EXAMPLE 4 Using $\nabla f(x, y)$ to Find a Directional Derivative

Find the directional derivative of

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

at $\left(-\frac{3}{4}, 0\right)$ in the direction from $P\left(-\frac{3}{4}, 0\right)$ to $Q(0, 1)$.

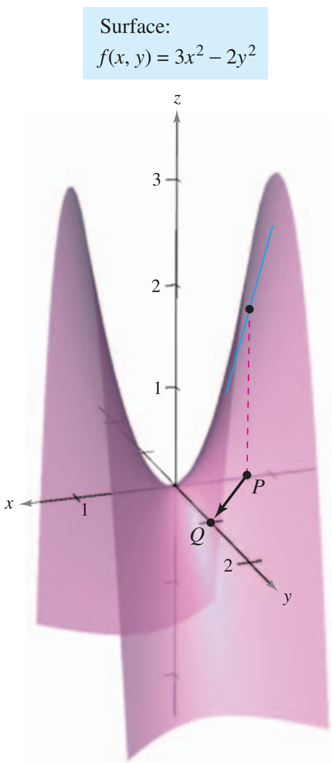


Figure 13.49

Rotatable Graph

Solution Because the partials of f are continuous, f is differentiable and you can apply Theorem 13.10. A vector in the specified direction is

$$\overrightarrow{PQ} = \mathbf{v} = \left(0 + \frac{3}{4}\right)\mathbf{i} + (1 - 0)\mathbf{j}$$

$$= \frac{3}{4}\mathbf{i} + \mathbf{j}$$

and a unit vector in this direction is

$$\mathbf{u} = \frac{\mathbf{v}}{\|\mathbf{v}\|} = \frac{3}{5}\mathbf{i} + \frac{4}{5}\mathbf{j}. \qquad \text{Unit vector in direction of } \overrightarrow{PQ}$$

Because $\nabla f(x, y) = f_x(x, y)\mathbf{i} + f_y(x, y)\mathbf{j} = 6x\mathbf{i} - 4y\mathbf{j}$, the gradient at $\left(-\frac{3}{4}, 0\right)$ is

$$\nabla f\left(-\frac{3}{4}, 0\right) = -\frac{9}{2}\mathbf{i} + 0\mathbf{j}. \qquad \text{Gradient at } \left(-\tfrac{3}{4}, 0\right)$$

Consequently, at $\left(-\frac{3}{4}, 0\right)$ the directional derivative is

$$D_{\mathbf{u}} f\left(-\frac{3}{4}, 0\right) = \nabla f\left(-\frac{3}{4}, 0\right) \cdot \mathbf{u}$$

$$= \left(-\frac{9}{2}\mathbf{i} + 0\mathbf{j}\right) \cdot \left(\frac{3}{5}\mathbf{i} + \frac{4}{5}\mathbf{j}\right)$$

$$= -\frac{27}{10}. \qquad \text{Directional derivative at } \left(-\tfrac{3}{4}, 0\right)$$

See Figure 13.49.

Try It   Exploration A

Applications of the Gradient

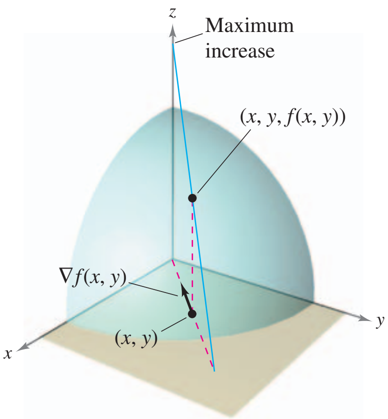
You have already seen that there are many directional derivatives at the point (x, y) on a surface. In many applications, you may want to know in which direction to move so that $f(x, y)$ increases most rapidly. This direction is called the direction of steepest ascent, and it is given by the gradient, as stated in the following theorem.

NOTE Part 2 of Theorem 13.11 says that at the point (x, y), f increases most rapidly in the direction of the gradient, $\nabla f(x, y)$.

THEOREM 13.11 Properties of the Gradient

Let f be differentiable at the point (x, y).

1. If $\nabla f(x, y) = \mathbf{0}$, then $D_{\mathbf{u}} f(x, y) = 0$ for all $\mathbf{u}$.
2. The direction of *maximum* increase of f is given by $\nabla f(x, y)$. The maximum value of $D_{\mathbf{u}} f(x, y)$ is $\|\nabla f(x, y)\|$.
3. The direction of *minimum* increase of f is given by $-\nabla f(x, y)$. The minimum value of $D_{\mathbf{u}} f(x, y)$ is $-\|\nabla f(x, y)\|$.



The gradient of f is a vector in the xy-plane that points in the direction of maximum increase on the surface given by $z = f(x, y)$.
Figure 13.50

Rotatable Graph

Proof If $\nabla f(x, y) = \mathbf{0}$, then for any direction (any $\mathbf{u}$), you have

$$\begin{aligned} D_{\mathbf{u}} f(x, y) &= \nabla f(x, y) \cdot \mathbf{u} \\ &= (0\mathbf{i} + 0\mathbf{j}) \cdot (\cos \theta \mathbf{i} + \sin \theta \mathbf{j}) \\ &= 0. \end{aligned}$$

If $\nabla f(x, y) \neq \mathbf{0}$, then let ϕ be the angle between $\nabla f(x, y)$ and a unit vector $\mathbf{u}$. Using the dot product, you can apply Theorem 11.5 to conclude that

$$\begin{aligned} D_{\mathbf{u}} f(x, y) &= \nabla f(x, y) \cdot \mathbf{u} \\ &= \|\nabla f(x, y)\| \, \|\mathbf{u}\| \cos \phi \\ &= \|\nabla f(x, y)\| \cos \phi \end{aligned}$$

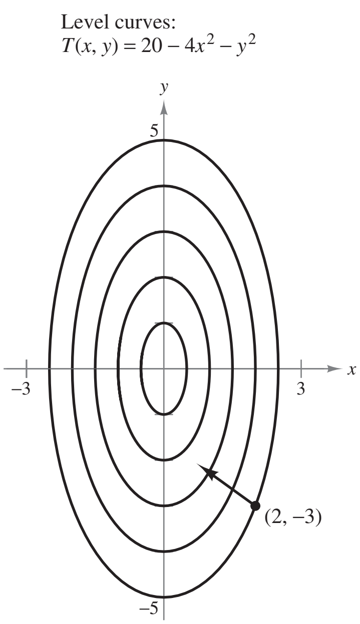
and it follows that the maximum value of $D_{\mathbf{u}} f(x, y)$ will occur when $\cos \phi = 1$. So, $\phi = 0$, and the maximum value for the directional derivative occurs when $\mathbf{u}$ has the same direction as $\nabla f(x, y)$. Moreover, this largest value for $D_{\mathbf{u}} f(x, y)$ is precisely

$$\|\nabla f(x, y)\| \cos \phi = \|\nabla f(x, y)\|.$$

Similarly, the minimum value of $D_{\mathbf{u}} f(x, y)$ can be obtained by letting $\phi = \pi$ so that $\mathbf{u}$ points in the direction opposite that of $\nabla f(x, y)$, as shown in Figure 13.50.

To visualize one of the properties of the gradient, imagine a skier coming down a mountainside. If $f(x, y)$ denotes the altitude of the skier, then $-\nabla f(x, y)$ indicates the *compass direction* the skier should take to ski the path of steepest descent. (Remember that the gradient indicates direction in the xy-plane and does not itself point up or down the mountainside.)

As another illustration of the gradient, consider the temperature $T(x, y)$ at any point (x, y) on a flat metal plate. In this case, $\nabla T(x, y)$ gives the direction of greatest temperature increase at the point (x, y), as illustrated in the next example.



The direction of most rapid increase in temperature at $(2, -3)$ is given by $-16\mathbf{i} + 6\mathbf{j}$.
Figure 13.51

EXAMPLE 5 Finding the Direction of Maximum Increase

The temperature in degrees Celsius on the surface of a metal plate is

$$T(x, y) = 20 - 4x^2 - y^2$$

where x and y are measured in centimeters. In what direction from $(2, -3)$ does the temperature increase most rapidly? What is this rate of increase?

Solution The gradient is

$$\begin{aligned} \nabla T(x, y) &= T_x(x, y)\mathbf{i} + T_y(x, y)\mathbf{j} \\ &= -8x\,\mathbf{i} - 2y\,\mathbf{j}. \end{aligned}$$

It follows that the direction of maximum increase is given by

$$\nabla T(2, -3) = -16\mathbf{i} + 6\mathbf{j}$$

as shown in Figure 13.51, and the rate of increase is

$$\begin{aligned} \|\nabla T(2, -3)\| &= \sqrt{256 + 36} \\ &= \sqrt{292} \\ &\approx 17.09^\circ \text{ per centimeter.} \end{aligned}$$

Try It Exploration A

The solution presented in Example 5 can be misleading. Although the gradient points in the direction of maximum temperature increase, it does not necessarily point toward the hottest spot on the plate. In other words, the gradient provides a local solution to finding an increase relative to the temperature at the point $(2, -3)$. *Once you leave that position, the direction of maximum increase may change.*

EXAMPLE 6 Finding the Path of a Heat-Seeking Particle

A heat-seeking particle is located at the point $(2, -3)$ on a metal plate whose temperature at (x, y) is

$$T(x, y) = 20 - 4x^2 - y^2.$$

Find the path of the particle as it continuously moves in the direction of maximum temperature increase.

Solution Let the path be represented by the position function

$$\mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j}.$$

A tangent vector at each point $(x(t), y(t))$ is given by

$$\mathbf{r}'(t) = \frac{dx}{dt}\mathbf{i} + \frac{dy}{dt}\mathbf{j}.$$

Because the particle seeks maximum temperature increase, the directions of $\mathbf{r}'(t)$ and $\nabla T(x, y) = -8x\mathbf{i} - 2y\mathbf{j}$ are the same at each point on the path. So,

$$-8x = k\frac{dx}{dt} \quad \text{and} \quad -2y = k\frac{dy}{dt}$$

where k depends on t. By solving each equation for dt/k and equating the results, you obtain

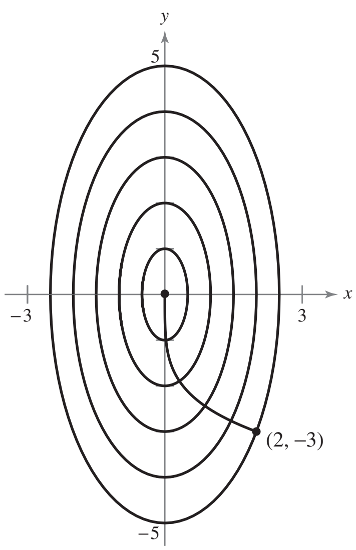
$$\frac{dx}{-8x} = \frac{dy}{-2y}.$$

The solution of this differential equation is $x = Cy^4$. Because the particle starts at the point $(2, -3)$, you can determine that $C = 2/81$. So, the path of the heat-seeking particle is

$$x = \frac{2}{81}y^4.$$

The path is shown in Figure 13.52.

Level curves:
$T(x, y) = 20 - 4x^2 - y^2$



Path followed by a heat-seeking particle
Figure 13.52

Try It | Exploration A

In Figure 13.52, the path of the particle (determined by the gradient at each point) appears to be orthogonal to each of the level curves. This becomes clear when you consider that the temperature $T(x, y)$ is constant along a given level curve. So, at any point (x, y) on the curve, the rate of change of T in the direction of a unit tangent vector $\mathbf{u}$ is 0, and you can write

$$\nabla f(x, y) \cdot \mathbf{u} = D_{\mathbf{u}}T(x, y) = 0. \qquad \mathbf{u} \text{ is a unit tangent vector.}$$

Because the dot product of $\nabla f(x, y)$ and $\mathbf{u}$ is 0, you can conclude that they must be orthogonal. This result is stated in the following theorem.

THEOREM 13.12 Gradient Is Normal to Level Curves

If f is differentiable at (x_0, y_0) and $\nabla f(x_0, y_0) \neq \mathbf{0}$, then $\nabla f(x_0, y_0)$ is normal to the level curve through (x_0, y_0).

EXAMPLE 7 Finding a Normal Vector to a Level Curve

Sketch the level curve corresponding to $c = 0$ for the function given by

$$f(x, y) = y - \sin x$$

and find a normal vector at several points on the curve.

Solution The level curve for $c = 0$ is given by

$$0 = y - \sin x$$
$$y = \sin x$$

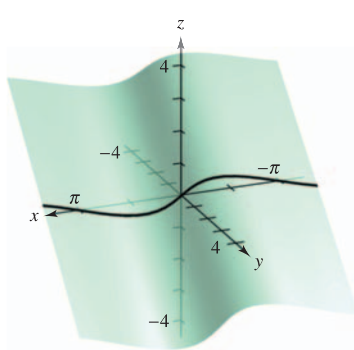
as shown in Figure 13.53(a). Because the gradient vector of f at (x, y) is

$$\begin{aligned}\nabla f(x, y) &= f_x(x, y)\mathbf{i} + f_y(x, y)\mathbf{j}\\ &= -\cos x\mathbf{i} + \mathbf{j}\end{aligned}$$

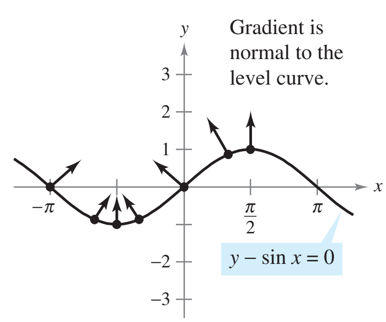
you can use Theorem 13.12 to conclude that $\nabla f(x, y)$ is normal to the level curve at the point (x, y). Some gradient vectors are

$$\nabla f(-\pi, 0) = \mathbf{i} + \mathbf{j}$$
$$\nabla f\left(-\frac{2\pi}{3}, -\frac{\sqrt{3}}{2}\right) = \frac{1}{2}\mathbf{i} + \mathbf{j}$$
$$\nabla f\left(-\frac{\pi}{2}, -1\right) = \mathbf{j}$$
$$\nabla f\left(-\frac{\pi}{3}, -\frac{\sqrt{3}}{2}\right) = -\frac{1}{2}\mathbf{i} + \mathbf{j}$$
$$\nabla f(0, 0) = -\mathbf{i} + \mathbf{j}$$
$$\nabla f\left(\frac{\pi}{3}, \frac{\sqrt{3}}{2}\right) = -\frac{1}{2}\mathbf{i} + \mathbf{j}$$
$$\nabla f\left(\frac{\pi}{2}, 1\right) = \mathbf{j}.$$

These are shown in Figure 13.53(b).



(a) The surface is given by $f(x, y) = y - \sin x$.



(b) The level curve is given by $f(x, y) = 0$.

Rotatable Graph

Figure 13.53

Try It

Exploration A

Functions of Three Variables

The definitions of the directional derivative and the gradient can be extended naturally to functions of three or more variables. As often happens, some of the geometric interpretation is lost in the generalization from functions of two variables to those of three variables. For example, you cannot interpret the directional derivative of a function of three variables to represent slope.

The definitions and properties of the directional derivative and the gradient of a function of three variables are given in the following summary.

Directional Derivative and Gradient for Three Variables

Let f be a function of x, y, and z, with continuous first partial derivatives. The **directional derivative of f** in the direction of a unit vector $\mathbf{u} = a\mathbf{i} + b\mathbf{j} + c\mathbf{k}$ is given by

$$D_{\mathbf{u}} f(x, y, z) = af_x(x, y, z) + bf_y(x, y, z) + cf_z(x, y, z).$$

The **gradient of f** is defined to be

$$\nabla f(x, y, z) = f_x(x, y, z)\mathbf{i} + f_y(x, y, z)\mathbf{j} + f_z(x, y, z)\mathbf{k}.$$

Properties of the gradient are as follows.

1. $D_{\mathbf{u}} f(x, y, z) = \nabla f(x, y, z) \cdot \mathbf{u}$
2. If $\nabla f(x, y, z) = \mathbf{0}$, then $D_{\mathbf{u}} f(x, y, z) = 0$ for all $\mathbf{u}$.
3. The direction of *maximum* increase of f is given by $\nabla f(x, y, z)$. The maximum value of $D_{\mathbf{u}} f(x, y, z)$ is

   $$\|\nabla f(x, y, z)\|. \qquad \text{Maximum value of } D_{\mathbf{u}} f(x, y, z)$$

4. The direction of *minimum* increase of f is given by $-\nabla f(x, y, z)$. The minimum value of $D_{\mathbf{u}} f(x, y, z)$ is

   $$-\|\nabla f(x, y, z)\|. \qquad \text{Minimum value of } D_{\mathbf{u}} f(x, y, z)$$

NOTE You can generalize Theorem 13.12 to functions of three variables. Under suitable hypotheses,

$$\nabla f(x_0, y_0, z_0)$$

is normal to the level surface through (x_0, y_0, z_0).

EXAMPLE 8 Finding the Gradient for a Function of Three Variables

Find $\nabla f(x, y, z)$ for the function given by

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

and find the direction of maximum increase of f at the point $(2, -1, 1)$.

Solution The gradient vector is given by

$$\begin{aligned}\nabla f(x, y, z) &= f_x(x, y, z)\mathbf{i} + f_y(x, y, z)\mathbf{j} + f_z(x, y, z)\mathbf{k} \\ &= 2x\,\mathbf{i} + 2y\,\mathbf{j} - 4\mathbf{k}.\end{aligned}$$

So, it follows that the direction of maximum increase at $(2, -1, 1)$ is

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

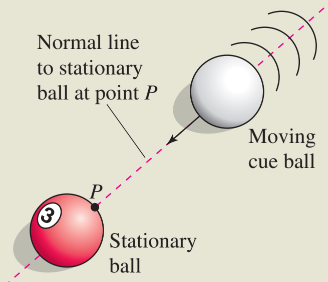
Try It   Exploration A

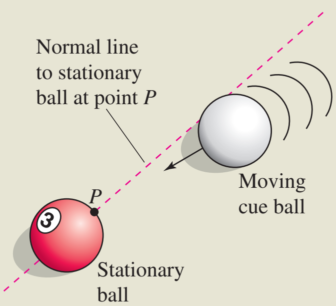
Section 13.7 Tangent Planes and Normal Lines

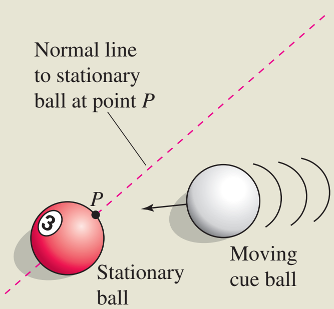
- **Find equations of tangent planes and normal lines to surfaces.**
- **Find the angle of inclination of a plane in space.**
- **Compare the gradients $\nabla f(x, y)$ and $\nabla F(x, y, z)$.**

EXPLORATION

Billiard Balls and Normal Lines
In each of the three figures below, the cue ball is about to strike a stationary ball at point P. Explain how you can use the normal line to the stationary ball at point P to describe the resulting motion of each of the two balls. Assuming that each cue ball has the same speed, which stationary ball will acquire the greatest speed? Which will acquire the least? Explain your reasoning.







Tangent Plane and Normal Line to a Surface

So far you have represented surfaces in space primarily by equations of the form

$$z = f(x, y). \qquad \text{Equation of a surface } S$$

In the development to follow, however, it is convenient to use the more general representation $F(x, y, z) = 0$. For a surface S given by $z = f(x, y)$, you can convert to the general form by defining F as

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

Because $f(x, y) - z = 0$, you can consider S to be the level surface of F given by

$$F(x, y, z) = 0. \qquad \text{Alternative equation of surface } S$$

EXAMPLE 1 Writing an Equation of a Surface

For the function given by

$$F(x, y, z) = x^2 + y^2 + z^2 - 4$$

describe the level surface given by $F(x, y, z) = 0$.

Solution The level surface given by $F(x, y, z) = 0$ can be written as

$$x^2 + y^2 + z^2 = 4$$

which is a sphere of radius 2 whose center is at the origin.

Try It Exploration A

You have seen many examples of the usefulness of normal lines in applications involving curves. Normal lines are equally important in analyzing surfaces and solids. For example, consider the collision of two billiard balls. When a stationary ball is struck at a point P on its surface, it moves along the **line of impact** determined by P and the center of the ball. The impact can occur in *two* ways. If the cue ball is moving along the line of impact, it stops dead and imparts all of its momentum to the stationary ball, as shown in Figure 13.54. If the cue ball is not moving along the line of impact, it is deflected to one side or the other and retains part of its momentum. That part of the momentum that is transferred to the stationary ball occurs along the line of impact, *regardless* of the direction of the cue ball, as shown in Figure 13.55. This line of impact is called the **normal line** to the surface of the ball at the point P.

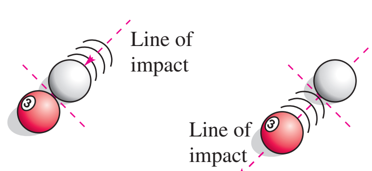


Figure 13.54

Animation

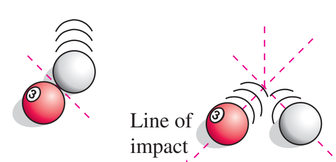


Figure 13.55

Animation

In the process of finding a normal line to a surface, you are also able to solve the problem of finding a **tangent plane** to the surface. Let S be a surface given by

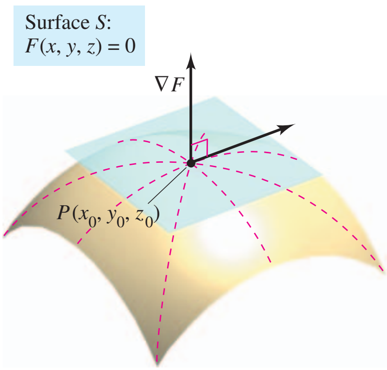
$$F(x, y, z) = 0$$

and let $P(x_0, y_0, z_0)$ be a point on S. Let C be a curve on S through P that is defined by the vector-valued function

$$\mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j} + z(t)\mathbf{k}.$$

Then, for all t,

$$F(x(t),\ y(t),\ z(t)) = 0.$$



Tangent plane to surface S at P
Figure 13.56

Rotatable Graph

If F is differentiable and $x'(t)$, $y'(t)$, and $z'(t)$ all exist, it follows from the Chain Rule that

$$\begin{aligned} 0 &= F'(t) \\ &= F_x(x, y, z)x'(t) + F_y(x, y, z)y'(t) + F_z(x, y, z)z'(t). \end{aligned}$$

At (x_0, y_0, z_0), the equivalent vector form is

$$0 = \underbrace{\nabla F(x_0, y_0, z_0)}_{\text{Gradient}} \cdot \underbrace{\mathbf{r}'(t_0)}_{\substack{\text{Tangent}\\ \text{vector}}}.$$

This result means that the gradient at P is orthogonal to the tangent vector of every curve on S through P. So, all tangent lines on S lie in a plane that is normal to $\nabla F(x_0, y_0, z_0)$ and contains P, as shown in Figure 13.56.

Definition of Tangent Plane and Normal Line

Let F be differentiable at the point $P(x_0, y_0, z_0)$ on the surface S given by $F(x, y, z) = 0$ such that $\nabla F(x_0, y_0, z_0) \neq \mathbf{0}$.

1. The plane through P that is normal to $\nabla F(x_0, y_0, z_0)$ is called the **tangent plane to S at P.**
2. The line through P having the direction of $\nabla F(x_0, y_0, z_0)$ is called the **normal line to S at P.**

NOTE In the remainder of this section, assume $\nabla F(x_0, y_0, z_0)$ to be nonzero unless stated otherwise.

To find an equation for the tangent plane to S at (x_0, y_0, z_0), let (x, y, z) be an arbitrary point in the tangent plane. Then the vector

$$\mathbf{v} = (x - x_0)\mathbf{i} + (y - y_0)\mathbf{j} + (z - z_0)\mathbf{k}$$

lies in the tangent plane. Because $\nabla F(x_0, y_0, z_0)$ is normal to the tangent plane at (x_0, y_0, z_0), it must be orthogonal to every vector in the tangent plane, and you have $\nabla F(x_0, y_0, z_0) \cdot \mathbf{v} = 0$, which leads to the following theorem.

THEOREM 13.13 Equation of Tangent Plane

If F is differentiable at (x_0, y_0, z_0), then an equation of the tangent plane to the surface given by $F(x, y, z) = 0$ at (x_0, y_0, z_0) is

$$F_x(x_0, y_0, z_0)(x - x_0) + F_y(x_0, y_0, z_0)(y - y_0) + F_z(x_0, y_0, z_0)(z - z_0) = 0.$$

EXAMPLE 2 Finding an Equation of a Tangent Plane

Find an equation of the tangent plane to the hyperboloid given by

$$z^2 - 2x^2 - 2y^2 = 12$$

at the point $(1, -1, 4)$.

Solution Begin by writing the equation of the surface as

$$z^2 - 2x^2 - 2y^2 - 12 = 0.$$

Then, considering

$$F(x, y, z) = z^2 - 2x^2 - 2y^2 - 12$$

you have

$$F_x(x, y, z) = -4x, \qquad F_y(x, y, z) = -4y, \quad \text{and} \quad F_z(x, y, z) = 2z.$$

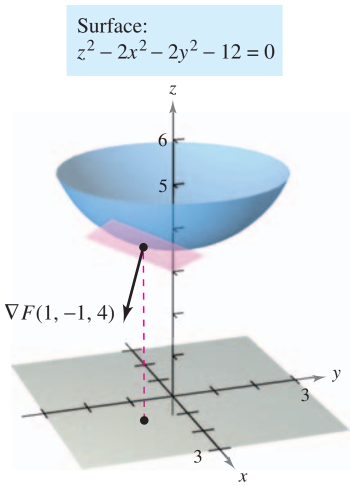
At the point $(1, -1, 4)$ the partial derivatives are

$$F_x(1, -1, 4) = -4, \qquad F_y(1, -1, 4) = 4, \quad \text{and} \quad F_z(1, -1, 4) = 8.$$

So, an equation of the tangent plane at $(1, -1, 4)$ is

$$\begin{aligned} -4(x - 1) + 4(y + 1) + 8(z - 4) &= 0 \\ -4x + 4 + 4y + 4 + 8z - 32 &= 0 \\ -4x + 4y + 8z - 24 &= 0 \\ x - y - 2z + 6 &= 0. \end{aligned}$$

Figure 13.57 shows a portion of the hyperboloid and tangent plane.

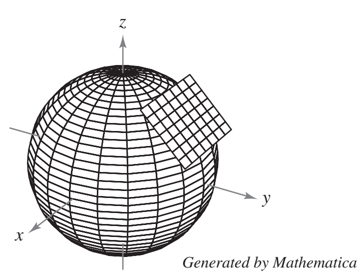


Tangent plane to surface
Figure 13.57

Rotatable Graph

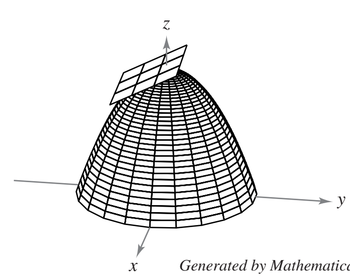
Try It Exploration A

TECHNOLOGY Some three-dimensional graphing utilities are capable of graphing tangent planes to surfaces. Two examples are shown below.



Sphere: $x^2 + y^2 + z^2 = 1$

Rotatable Graph



Paraboloid: $z = 2 - x^2 - y^2$

Rotatable Graph

To find the equation of the tangent plane at a point on a surface given by $z = f(x, y)$, you can define the function F by

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

Then S is given by the level surface $F(x, y, z) = 0$, and by Theorem 13.13 an equation of the tangent plane to S at the point (x_0, y_0, z_0) is

$$f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0) - (z - z_0) = 0.$$

EXAMPLE 3 Finding an Equation of the Tangent Plane

Find the equation of the tangent plane to the paraboloid

$$z = 1 - \frac{1}{10}(x^2 + 4y^2)$$

at the point $\left(1, 1, \frac{1}{2}\right)$.

Solution From $z = f(x, y) = 1 - \frac{1}{10}(x^2 + 4y^2)$, you obtain

$$f_x(x, y) = -\frac{x}{5} \quad \Rightarrow \quad f_x(1, 1) = -\frac{1}{5}$$

and

$$f_y(x, y) = -\frac{4y}{5} \quad \Rightarrow \quad f_y(1, 1) = -\frac{4}{5}.$$

So, an equation of the tangent plane at $\left(1, 1, \frac{1}{2}\right)$ is

$$f_x(1, 1)(x - 1) + f_y(1, 1)(y - 1) - \left(z - \frac{1}{2}\right) = 0$$

$$-\frac{1}{5}(x - 1) - \frac{4}{5}(y - 1) - \left(z - \frac{1}{2}\right) = 0$$

$$-\frac{1}{5}x - \frac{4}{5}y - z + \frac{3}{2} = 0.$$

This tangent plane is shown in Figure 13.58.

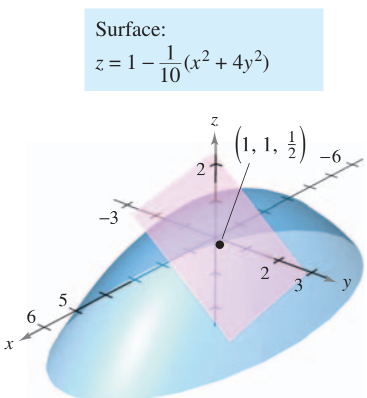


Figure 13.58

Rotatable Graph

Try It | Exploration A

The gradient $\nabla F(x, y, z)$ gives a convenient way to find equations of normal lines, as shown in Example 4.

EXAMPLE 4 Finding an Equation of a Normal Line to a Surface

Find a set of symmetric equations for the normal line to the surface given by $xyz = 12$ at the point $(2, -2, -3)$.

Solution Begin by letting

$$F(x, y, z) = xyz - 12.$$

Then, the gradient is given by

$$\begin{aligned}\nabla F(x, y, z) &= F_x(x, y, z)\mathbf{i} + F_y(x, y, z)\mathbf{j} + F_z(x, y, z)\mathbf{k} \\ &= yz\mathbf{i} + xz\mathbf{j} + xy\mathbf{k}\end{aligned}$$

and at the point $(2, -2, -3)$ you have

$$\begin{aligned}\nabla F(2, -2, -3) &= (-2)(-3)\mathbf{i} + (2)(-3)\mathbf{j} + (2)(-2)\mathbf{k} \\ &= 6\mathbf{i} - 6\mathbf{j} - 4\mathbf{k}.\end{aligned}$$

The normal line at $(2, -2, -3)$ has direction numbers 6, -6, and -4, and the corresponding set of symmetric equations is

$$\frac{x - 2}{6} = \frac{y + 2}{-6} = \frac{z + 3}{-4}.$$

See Figure 13.59.

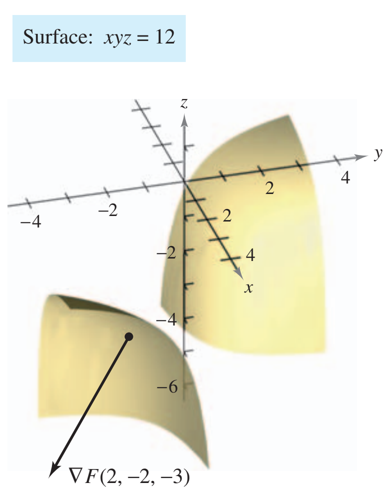


Figure 13.59

Rotatable Graph

Try It | Exploration A | Exploration B | Open Exploration

Knowing that the gradient $\nabla F(x, y, z)$ is normal to the surface given by $F(x, y, z) = 0$ allows you to solve a variety of problems dealing with surfaces and curves in space.

EXAMPLE 5 Finding the Equation of a Tangent Line to a Curve

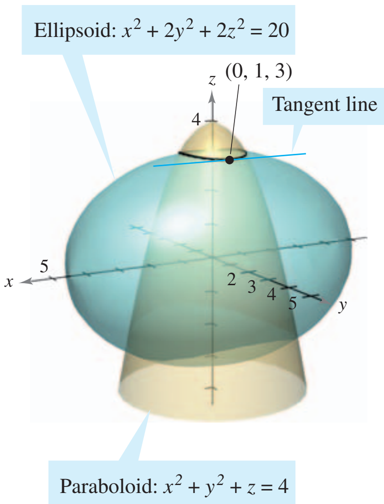


Figure 13.60

Rotatable Graph

Describe the tangent line to the curve of intersection of the surfaces

$$x^2 + 2y^2 + 2z^2 = 20 \qquad \text{Ellipsoid}$$
$$x^2 + y^2 + z = 4 \qquad \text{Paraboloid}$$

at the point (0, 1, 3), as shown in Figure 13.60.

Solution Begin by finding the gradients to both surfaces at the point (0, 1, 3).

Ellipsoid	*Paraboloid*
$F(x, y, z) = x^2 + 2y^2 + 2z^2 - 20$	$G(x, y, z) = x^2 + y^2 + z - 4$
$\nabla F(x, y, z) = 2x\mathbf{i} + 4y\mathbf{j} + 4z\mathbf{k}$	$\nabla G(x, y, z) = 2x\mathbf{i} + 2y\mathbf{j} + \mathbf{k}$
$\nabla F(0, 1, 3) = 4\mathbf{j} + 12\mathbf{k}$	$\nabla G(0, 1, 3) = 2\mathbf{j} + \mathbf{k}$

The cross product of these two gradients is a vector that is tangent to both surfaces at the point (0, 1, 3).

$$\nabla F(0, 1, 3) \times \nabla G(0, 1, 3) = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 0 & 4 & 12 \\ 0 & 2 & 1 \end{vmatrix} = -20\mathbf{i}.$$

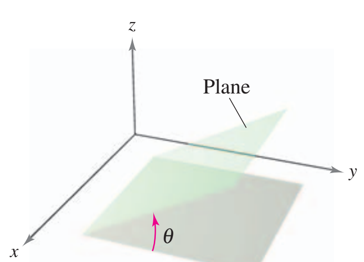
So, the tangent line to the curve of intersection of the two surfaces at the point (0, 1, 3) is a line that is parallel to the x-axis and passes through the point (0, 1, 3).

Try It Exploration A

The Angle of Inclination of a Plane

Another use of the gradient $\nabla F(x, y, z)$ is to determine the angle of inclination of the tangent plane to a surface. The **angle of inclination** of a plane is defined to be the angle θ $(0 \le \theta \le \pi/2)$ between the given plane and the xy-plane, as shown in Figure 13.61. (The angle of inclination of a horizontal plane is defined to be zero.) Because the vector $\mathbf{k}$ is normal to the xy-plane, you can use the formula for the cosine of the angle between two planes (given in Section 11.5) to conclude that the angle of inclination of a plane with normal vector $\mathbf{n}$ is given by

$$\cos\theta = \frac{|\mathbf{n} \cdot \mathbf{k}|}{\|\mathbf{n}\|\,\|\mathbf{k}\|} = \frac{|\mathbf{n} \cdot \mathbf{k}|}{\|\mathbf{n}\|}. \qquad \text{Angle of inclination of a plane}$$



The angle of inclination
Figure 13.61

Rotatable Graph

EXAMPLE 6 Finding the Angle of Inclination of a Tangent Plane

Find the angle of inclination of the tangent plane to the ellipsoid given by

$$\frac{x^2}{12} + \frac{y^2}{12} + \frac{z^2}{3} = 1$$

at the point $(2, 2, 1)$.

Solution If you let

$$F(x, y, z) = \frac{x^2}{12} + \frac{y^2}{12} + \frac{z^2}{3} - 1$$

the gradient of F at the point $(2, 2, 1)$ is given by

$$\nabla F(x, y, z) = \frac{x}{6}\mathbf{i} + \frac{y}{6}\mathbf{j} + \frac{2z}{3}\mathbf{k}$$

$$\nabla F(2, 2, 1) = \frac{1}{3}\mathbf{i} + \frac{1}{3}\mathbf{j} + \frac{2}{3}\mathbf{k}.$$

Because $\nabla F(2, 2, 1)$ is normal to the tangent plane and $\mathbf{k}$ is normal to the xy-plane, it follows that the angle of inclination of the tangent plane is given by

$$\cos\theta = \frac{|\nabla F(2, 2, 1) \cdot \mathbf{k}|}{\|\nabla F(2, 2, 1)\|} = \frac{2/3}{\sqrt{(1/3)^2 + (1/3)^2 + (2/3)^2}} = \sqrt{\frac{2}{3}}$$

which implies that

$$\theta = \arccos\sqrt{\frac{2}{3}} \approx 35.3^\circ,$$

as shown in Figure 13.62.

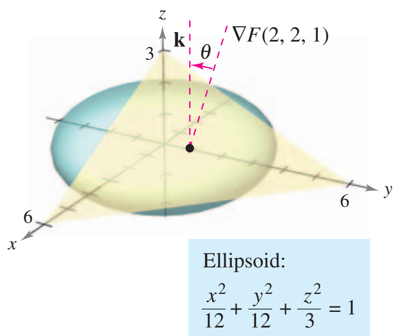


Figure 13.62

Rotatable Graph

Try It

Exploration A

NOTE A special case of the procedure shown in Example 6 is worth noting. The angle of inclination θ of the tangent plane to the surface $z = f(x, y)$ at (x_0, y_0, z_0) is given by

$$\cos\theta = \frac{1}{\sqrt{[f_x(x_0, y_0)]^2 + [f_y(x_0, y_0)]^2 + 1}}.$$

Alternative formula for angle of inclination (See Exercise 64.)

A Comparison of the Gradients $\nabla f(x, y)$ and $\nabla F(x, y, z)$

This section concludes with a comparison of the gradients $\nabla f(x, y)$ and $\nabla F(x, y, z)$. In the preceding section, you saw that the gradient of a function f of two variables is normal to the level curves of f. Specifically, Theorem 13.12 states that if f is differentiable at (x_0, y_0) and $\nabla f(x_0, y_0) \neq \mathbf{0}$, then $\nabla f(x_0, y_0)$ is normal to the level curve through (x_0, y_0). Having developed normal lines to surfaces, you can now extend this result to a function of three variables. The proof of Theorem 13.14 is left as an exercise (see Exercise 63).

THEOREM 13.14 Gradient Is Normal to Level Surfaces

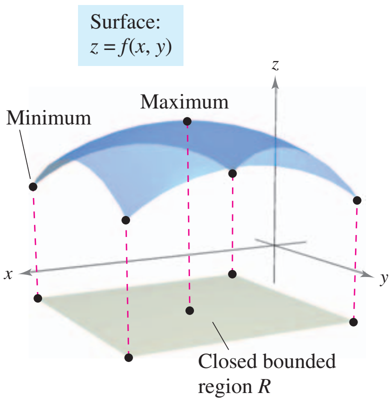
If F is differentiable at (x_0, y_0, z_0) and $\nabla F(x_0, y_0, z_0) \neq \mathbf{0}$, then $\nabla F(x_0, y_0, z_0)$ is normal to the level surface through (x_0, y_0, z_0).

When working with the gradients $\nabla f(x, y)$ and $\nabla F(x, y, z)$, be sure you remember that $\nabla f(x, y)$ is a vector in the xy-plane and $\nabla F(x, y, z)$ is a vector in space.

Section 13.8 Extrema of Functions of Two Variables

- **Find absolute and relative extrema of a function of two variables.**
- **Use the Second Partials Test to find relative extrema of a function of two variables.**

Absolute Extrema and Relative Extrema



R contains point(s) at which $f(x, y)$ is a minimum and point(s) at which $f(x, y)$ is a maximum.
Figure 13.63

Rotatable Graph

In Chapter 3, you studied techniques for finding the extreme values of a function of a single variable. In this section, you will extend these techniques to functions of two variables. For example, in Theorem 13.15 the Extreme Value Theorem for a function of a single variable is extended to a function of two variables.

Consider the continuous function f of two variables, defined on a closed bounded region R. The values $f(a, b)$ and $f(c, d)$ such that

$$f(a, b) \le f(x, y) \le f(c, d) \qquad (a, b) \text{ and } (c, d) \text{ are in } R.$$

for all (x, y) in R are called the **minimum** and **maximum** of f in the region R, as shown in Figure 13.63. Recall from Section 13.2 that a region in the plane is *closed* if it contains all of its boundary points. The Extreme Value Theorem deals with a region in the plane that is both closed and *bounded*. A region in the plane is called **bounded** if it is a subregion of a closed disk in the plane.

THEOREM 13.15 Extreme Value Theorem

Let f be a continuous function of two variables x and y defined on a closed bounded region R in the xy-plane.

1. There is at least one point in R where f takes on a minimum value.
2. There is at least one point in R where f takes on a maximum value.

A minimum is also called an **absolute minimum** and a maximum is also called an **absolute maximum.** As in single-variable calculus, there is a distinction made between absolute extrema and **relative extrema.**

Definition of Relative Extrema

Let f be a function defined on a region R containing (x_0, y_0).

1. The function f has a **relative minimum** at (x_0, y_0) if

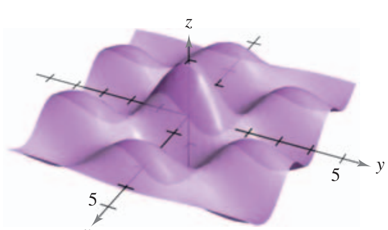
$$f(x, y) \ge f(x_0, y_0)$$

for all (x, y) in an *open* disk containing (x_0, y_0).

2. The function f has a **relative maximum** at (x_0, y_0) if

$$f(x, y) \le f(x_0, y_0)$$

for all (x, y) in an *open* disk containing (x_0, y_0).



Relative extrema
Figure 13.64

Rotatable Graph

To say that f has a relative maximum at (x_0, y_0) means that the point (x_0, y_0, z_0) is at least as high as all nearby points on the graph of $z = f(x, y)$. Similarly, f has a relative minimum at (x_0, y_0) if (x_0, y_0, z_0) is at least as low as all nearby points on the graph. (See Figure 13.64.)

KARL WEIERSTRASS (1815–1897)

Although the Extreme Value Theorem had been used by earlier mathematicians, the first to provide a rigorous proof was the German mathematician Karl Weierstrass. Weierstrass also provided rigorous justifications for many other mathematical results already in common use. We are indebted to him for much of the logical foundation on which modern calculus is built.

MathBio

To locate relative extrema of f, you can investigate the points at which the gradient of f is **0** or the points at which one of the partial derivatives does not exist. Such points are called **critical points** of f.

Definition of Critical Point

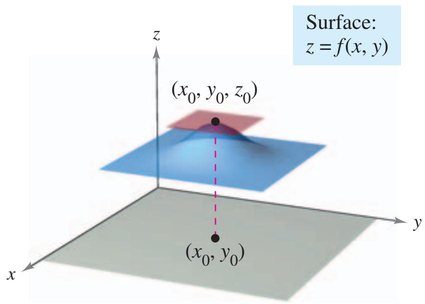
Let f be defined on an open region R containing (x_0, y_0). The point (x_0, y_0) is a **critical point** of f if one of the following is true.

1. $f_x(x_0, y_0) = 0$ and $f_y(x_0, y_0) = 0$
2. $f_x(x_0, y_0)$ or $f_y(x_0, y_0)$ does not exist.

Recall from Theorem 13.11 that if f is differentiable and

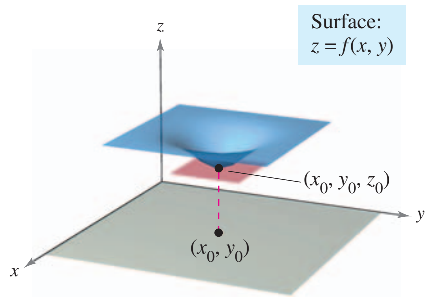
$$\begin{aligned}\nabla f(x_0, y_0) &= f_x(x_0, y_0)\mathbf{i} + f_y(x_0, y_0)\mathbf{j}\\ &= 0\mathbf{i} + 0\mathbf{j}\end{aligned}$$

then every directional derivative at (x_0, y_0) must be 0. This implies that the function has a horizontal tangent plane at the point (x_0, y_0), as shown in Figure 13.65. It appears that such a point is a likely location of a relative extremum. This is confirmed by Theorem 13.16.



Relative maximum

Rotatable Graph



Relative minimum

Rotatable Graph

Figure 13.65

THEOREM 13.16 Relative Extrema Occur Only at Critical Points

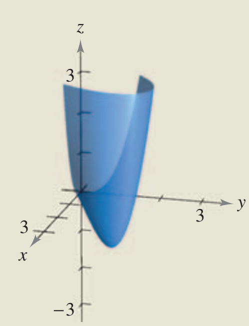
If f has a relative extremum at (x_0, y_0) on an open region R, then (x_0, y_0) is a critical point of f.

EXPLORATION

Use a graphing utility to graph

$$z = x^3 - 3xy + y^3$$

using the bounds $0 \le x \le 3$, $0 \le y \le 3$, and $-3 \le z \le 3$. This view makes it appear as though the surface has an absolute minimum. But does it?



Rotatable Graph

EXAMPLE 1 Finding a Relative Extremum

Determine the relative extrema of

$$f(x, y) = 2x^2 + y^2 + 8x - 6y + 20.$$

Solution Begin by finding the critical points of f. Because

$$f_x(x, y) = 4x + 8 \qquad \text{Partial with respect to } x$$

and

$$f_y(x, y) = 2y - 6 \qquad \text{Partial with respect to } y$$

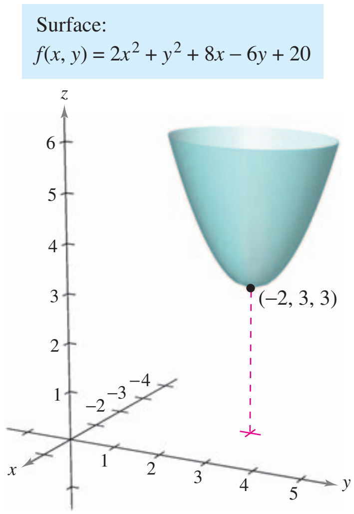
are defined for all x and y, the only critical points are those for which both first partial derivatives are 0. To locate these points, let $f_x(x, y)$ and $f_y(x, y)$ be 0, and solve the equations

$$4x + 8 = 0 \quad \text{and} \quad 2y - 6 = 0$$

to obtain the critical point $(-2, 3)$. By completing the square, you can conclude that for all $(x, y) \neq (-2, 3)$

$$f(x, y) = 2(x + 2)^2 + (y - 3)^2 + 3 > 3.$$

So, a relative *minimum* of f occurs at $(-2, 3)$. The value of the relative minimum is $f(-2, 3) = 3$, as shown in Figure 13.66.



The function $z = f(x, y)$ has a relative minimum at $(-2, 3)$.
Figure 13.66

Rotatable Graph

Example 1 shows a relative minimum occurring at one type of critical point—the type for which both $f_x(x, y)$ and $f_y(x, y)$ are 0. The next example concerns a relative maximum that occurs at the other type of critical point—the type for which either $f_x(x, y)$ or $f_y(x, y)$ does not exist.

Try It | Exploration A | Exploration B | Open Exploration

EXAMPLE 2 Finding a Relative Extremum

Determine the relative extrema of $f(x, y) = 1 - (x^2 + y^2)^{1/3}$.

Solution Because

$$f_x(x, y) = -\frac{2x}{3(x^2 + y^2)^{2/3}} \qquad \text{Partial with respect to } x$$

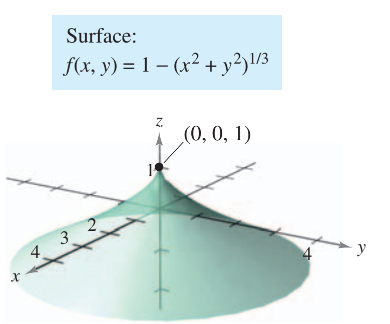
and

$$f_y(x, y) = -\frac{2y}{3(x^2 + y^2)^{2/3}} \qquad \text{Partial with respect to } y$$

it follows that both partial derivatives exist for all points in the xy-plane except for $(0, 0)$. Moreover, because the partial derivatives cannot both be 0 unless both x and y are 0, you can conclude that $(0, 0)$ is the only critical point. In Figure 13.67, note that $f(0, 0)$ is 1. For all other (x, y) it is clear that

$$f(x, y) = 1 - (x^2 + y^2)^{1/3} < 1.$$

So, f has a relative *maximum* at $(0, 0)$.

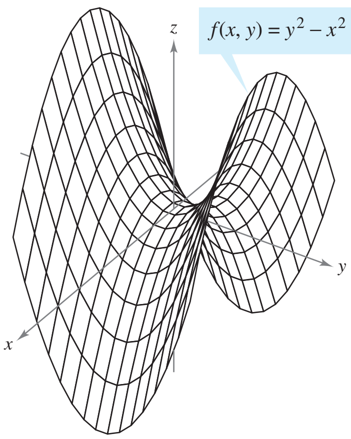


$f_x(x, y)$ and $f_y(x, y)$ are undefined at $(0, 0)$.
Figure 13.67

Rotatable Graph

Try It | Exploration A

NOTE In Example 2, $f_x(x, y) = 0$ for every point on the y-axis other than $(0, 0)$. However, because $f_y(x, y)$ is nonzero, these are not critical points. Remember that *one* of the partials must not exist or *both* must be 0 in order to yield a critical point.



Saddle point at $(0, 0, 0)$:
$f_x(0, 0) = f_y(0, 0) = 0$

Figure 13.68

Rotatable Graph

The Second Partials Test

Theorem 13.16 tells you that to find relative extrema you need only examine values of $f(x, y)$ at critical points. However, as is true for a function of one variable, the critical points of a function of two variables do not always yield relative maxima or minima. Some critical points yield **saddle points,** which are neither relative maxima nor relative minima.

As an example of a critical point that does not yield a relative extremum, consider the surface given by

$$f(x, y) = y^2 - x^2 \qquad \text{Hyperbolic paraboloid}$$

as shown in Figure 13.68. At the point $(0, 0)$, both partial derivatives are 0. The function f does not, however, have a relative extremum at this point because in any open disk centered at $(0, 0)$ the function takes on both negative values (along the x-axis) *and* positive values (along the y-axis). So, the point $(0, 0, 0)$ is a saddle point of the surface. (The term "saddle point" comes from the fact that the surface shown in Figure 13.68 resembles a saddle.)

For the functions in Examples 1 and 2, it was relatively easy to determine the relative extrema, because each function was either given, or able to be written, in completed square form. For more complicated functions, algebraic arguments are less convenient and it is better to rely on the analytic means presented in the following Second Partials Test. This is the two-variable counterpart of the Second Derivative Test for functions of one variable. The proof of this theorem is best left to a course in advanced calculus.

THEOREM 13.17 Second Partials Test

Let f have continuous second partial derivatives on an open region containing a point (a, b) for which

$$f_x(a, b) = 0 \quad \text{and} \quad f_y(a, b) = 0.$$

To test for relative extrema of f, consider the quantity

$$d = f_{xx}(a, b)f_{yy}(a, b) - [f_{xy}(a, b)]^2.$$

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 $(a, b, f(a, b))$ is a **saddle point.**
4. The test is inconclusive if $d = 0$.

NOTE If $d > 0$, then $f_{xx}(a, b)$ and $f_{yy}(a, b)$ must have the same sign. This means that $f_{xx}(a, b)$ can be replaced by $f_{yy}(a, b)$ in the first two parts of the test.

A convenient device for remembering the formula for d in the Second Partials Test is given by the 2×2 determinant

$$d = \begin{vmatrix} f_{xx}(a, b) & f_{xy}(a, b) \\ f_{yx}(a, b) & f_{yy}(a, b) \end{vmatrix}$$

where $f_{xy}(a, b) = f_{yx}(a, b)$ by Theorem 13.3.

EXAMPLE 3 Using the Second Partials Test

Find the relative extrema of $f(x, y) = -x^3 + 4xy - 2y^2 + 1$.

Solution Begin by finding the critical points of f. Because

$$f_x(x, y) = -3x^2 + 4y \quad \text{and} \quad f_y(x, y) = 4x - 4y$$

exist for all x and y, the only critical points are those for which both first partial derivatives are 0. To locate these points, let $f_x(x, y)$ and $f_y(x, y)$ be 0 to obtain $-3x^2 + 4y = 0$ and $4x - 4y = 0$. From the second equation you know that $x = y$, and, by substitution into the first equation, you obtain two solutions: $y = x = 0$ and $y = x = \frac{4}{3}$. Because

$$f_{xx}(x, y) = -6x, \quad f_{yy}(x, y) = -4, \quad \text{and} \quad f_{xy}(x, y) = 4$$

it follows that, for the critical point $(0, 0)$,

$$d = f_{xx}(0, 0)f_{yy}(0, 0) - [f_{xy}(0, 0)]^2 = 0 - 16 < 0$$

and, by the Second Partials Test, you can conclude that $(0, 0, 1)$ is a saddle point of f. Furthermore, for the critical point $\left(\frac{4}{3}, \frac{4}{3}\right)$,

$$\begin{aligned} d &= f_{xx}\left(\tfrac{4}{3}, \tfrac{4}{3}\right) f_{yy}\left(\tfrac{4}{3}, \tfrac{4}{3}\right) - \left[f_{xy}\left(\tfrac{4}{3}, \tfrac{4}{3}\right)\right]^2 \\ &= -8(-4) - 16 \\ &= 16 \\ &> 0 \end{aligned}$$

and because $f_{xx}\left(\frac{4}{3}, \frac{4}{3}\right) = -8 < 0$ you can conclude that f has a relative maximum at $\left(\frac{4}{3}, \frac{4}{3}\right)$, as shown in Figure 13.69.

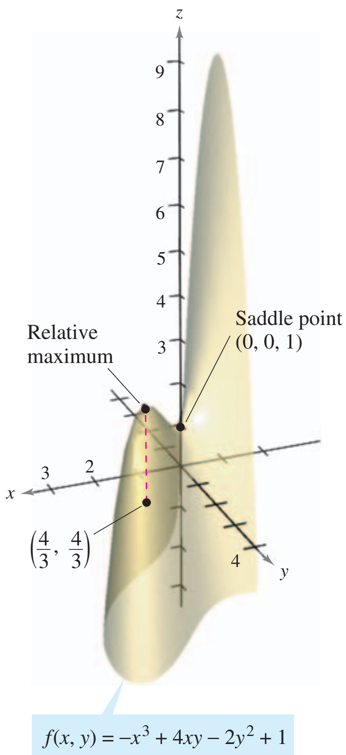


Figure 13.69

Rotatable Graph

Try It

Exploration A

The Second Partials Test can fail to find relative extrema in two ways. If either of the first partial derivatives does not exist, you cannot use the test. Also, if

$$d = f_{xx}(a, b)f_{yy}(a, b) - [f_{xy}(a, b)]^2 = 0$$

the test fails. In such cases, you can try a sketch or some other approach, as demonstrated in the next example.

EXAMPLE 4 Failure of the Second Partials Test

Find the relative extrema of $f(x, y) = x^2y^2$.

Solution Because $f_x(x, y) = 2xy^2$ and $f_y(x, y) = 2x^2y$, you know that both partial derivatives are 0 if $x = 0$ or $y = 0$. That is, every point along the x- or y-axis is a critical point. Moreover, because

$$f_{xx}(x, y) = 2y^2, \qquad f_{yy}(x, y) = 2x^2, \qquad \text{and} \qquad f_{xy}(x, y) = 4xy$$

you know that if either $x = 0$ or $y = 0$, then

$$\begin{aligned} d &= f_{xx}(x, y)f_{yy}(x, y) - [f_{xy}(x, y)]^2 \\ &= 4x^2y^2 - 16x^2y^2 = -12x^2y^2 = 0. \end{aligned}$$

So, the Second Partials Test fails. However, because $f(x, y) = 0$ for every point along the x- or y-axis and $f(x, y) = x^2y^2 > 0$ for all other points, you can conclude that each of these critical points yields an absolute minimum, as shown in Figure 13.70.

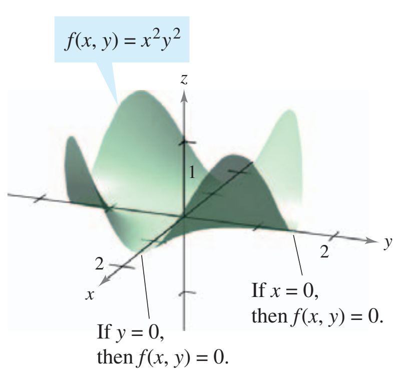


Figure 13.70

Rotatable Graph

Try It

Exploration A

Absolute extrema of a function can occur in two ways. First, some relative extrema also happen to be absolute extrema. For instance, in Example 1, $f(-2, 3)$ is an absolute minimum of the function. (On the other hand, the relative maximum found in Example 3 is not an absolute maximum of the function.) Second, absolute extrema can occur at a boundary point of the domain. This is illustrated in Example 5.

EXAMPLE 5 Finding Absolute Extrema

Find the absolute extrema of the function

$$f(x, y) = \sin xy$$

on the closed region given by $0 \le x \le \pi$ and $0 \le y \le 1$.

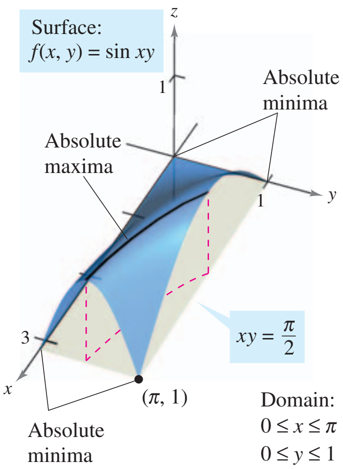


Figure 13.71

Rotatable Graph

Solution From the partial derivatives

$$f_x(x, y) = y \cos xy \quad \text{and} \quad f_y(x, y) = x \cos xy$$

you can see that each point lying on the hyperbola given by $xy = \pi/2$ is a critical point. These points each yield the value

$$f(x, y) = \sin \frac{\pi}{2} = 1$$

which you know is the absolute maximum, as shown in Figure 13.71. The only other critical point of f *lying in the given region* is $(0, 0)$. It yields an absolute minimum of 0, because

$$0 \le xy \le \pi$$

implies that

$$0 \le \sin xy \le 1.$$

To locate other absolute extrema, you should consider the four boundaries of the region formed by taking traces with the vertical planes $x = 0$, $x = \pi$, $y = 0$, and $y = 1$. In doing this, you will find that $\sin xy = 0$ at all points on the x-axis, at all points on the y-axis, and at the point $(\pi, 1)$. Each of these points yields an absolute minimum for the surface, as shown in Figure 13.71.

Try It Exploration A

The concepts of relative extrema and critical points can be extended to functions of three or more variables. If all first partial derivatives of

$$w = f(x_1, x_2, x_3, \ldots, x_n)$$

exist, it can be shown that a relative maximum or minimum can occur at $(x_1, x_2, x_3, \ldots, x_n)$ only if every first partial derivative is 0 at that point. This means that the critical points are obtained by solving the following system of equations.

$$\begin{aligned} f_{x_1}(x_1, x_2, x_3, \ldots, x_n) &= 0 \\ f_{x_2}(x_1, x_2, x_3, \ldots, x_n) &= 0 \\ &\vdots \\ f_{x_n}(x_1, x_2, x_3, \ldots, x_n) &= 0 \end{aligned}$$

The extension of Theorem 13.17 to three or more variables is also possible, although you will not consider such an extension in this text.

Section 13.9 Applications of Extrema of Functions of Two Variables

- **Solve optimization problems involving functions of several variables.**
- **Use the method of least squares.**

Applied Optimization Problems

In this section, you will survey a few of the many applications of extrema of functions of two (or more) variables.

EXAMPLE 1 Finding Maximum Volume

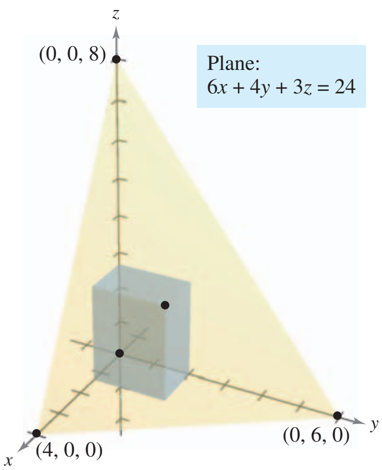


Figure 13.72

Rotatable Graph

A rectangular box is resting on the xy-plane with one vertex at the origin. The opposite vertex lies in the plane

$$6x + 4y + 3z = 24$$

as shown in Figure 13.72. Find the maximum volume of such a box.

Solution Let x, y, and z represent the length, width, and height of the box. Because one vertex of the box lies in the plane $6x + 4y + 3z = 24$, you know that $z = \frac{1}{3}(24 - 6x - 4y)$, and you can write the volume xyz of the box as a function of two variables.

$$\begin{aligned} V(x, y) &= (x)(y)\left[\tfrac{1}{3}(24 - 6x - 4y)\right] \\ &= \tfrac{1}{3}(24xy - 6x^2y - 4xy^2) \end{aligned}$$

By setting the first partial derivatives equal to 0

$$V_x(x, y) = \tfrac{1}{3}(24y - 12xy - 4y^2) = \frac{y}{3}(24 - 12x - 4y) = 0$$

$$V_y(x, y) = \tfrac{1}{3}(24x - 6x^2 - 8xy) = \frac{x}{3}(24 - 6x - 8y) = 0$$

you obtain the critical points $(0, 0)$ and $\left(\frac{4}{3}, 2\right)$. At $(0, 0)$ the volume is 0, so that point does not yield a maximum volume. At the point $\left(\frac{4}{3}, 2\right)$, you can apply the Second Partials Test.

$$\begin{aligned} V_{xx}(x, y) &= -4y \\ V_{yy}(x, y) &= \frac{-8x}{3} \\ V_{xy}(x, y) &= \tfrac{1}{3}(24 - 12x - 8y) \end{aligned}$$

Because

$$V_{xx}\left(\tfrac{4}{3}, 2\right)V_{yy}\left(\tfrac{4}{3}, 2\right) - \left[V_{xy}\left(\tfrac{4}{3}, 2\right)\right]^2 = (-8)\left(-\tfrac{32}{9}\right) - \left(-\tfrac{8}{3}\right)^2 = \tfrac{64}{3} > 0$$

and

$$V_{xx}\left(\tfrac{4}{3}, 2\right) = -8 < 0$$

you can conclude from the Second Partials Test that the maximum volume is

$$\begin{aligned} V\left(\tfrac{4}{3}, 2\right) &= \tfrac{1}{3}\left[24\left(\tfrac{4}{3}\right)(2) - 6\left(\tfrac{4}{3}\right)^2(2) - 4\left(\tfrac{4}{3}\right)(2^2)\right] \\ &= \tfrac{64}{9} \text{ cubic units.} \end{aligned}$$

Note that the volume is 0 at the boundary points of the triangular domain of V.

Try It | Exploration A | Open Exploration

NOTE In many applied problems, the domain of the function to be optimized is a closed bounded region. To find minimum or maximum points, you must not only test critical points, but also consider the values of the function at points on the boundary.

Applications of extrema in economics and business often involve more than one independent variable. For instance, a company may produce several models of one type of product. The price per unit and profit per unit are usually different for each model. Moreover, the demand for each model is often a function of the prices of the other models (as well as its own price). The next example illustrates an application involving two products.

EXAMPLE 2 Finding the Maximum Profit

An electronics manufacturer determines that the profit P (in dollars) obtained by producing x units of a DVD player and y units of a DVD recorder is approximated by the model

$$P(x, y) = 8x + 10y - (0.001)(x^2 + xy + y^2) - 10{,}000.$$

Find the production level that produces a maximum profit. What is the maximum profit?

Solution

The partial derivatives of the profit function are

$$P_x(x, y) = 8 - (0.001)(2x + y) \qquad \text{and} \qquad P_y(x, y) = 10 - (0.001)(x + 2y).$$

FOR FURTHER INFORMATION For more information on the use of mathematics in economics, see the article "Mathematical Methods of Economics" by Joel Franklin in *The American Mathematical Monthly*.

MathArticle

By setting these partial derivatives equal to 0, you obtain the following system of equations.

$$\begin{aligned} 8 - (0.001)(2x + y) &= 0 \\ 10 - (0.001)(x + 2y) &= 0 \end{aligned}$$

After simplifying, this system of linear equations can be written as

$$\begin{aligned} 2x + y &= 8000 \\ x + 2y &= 10{,}000. \end{aligned}$$

Solving this system produces $x = 2000$ and $y = 4000$. The second partial derivatives of P are

$$\begin{aligned} P_{xx}(2000, 4000) &= -0.002 \\ P_{yy}(2000, 4000) &= -0.002 \\ P_{xy}(2000, 4000) &= -0.001. \end{aligned}$$

Because $P_{xx} < 0$ and

$$P_{xx}(2000, 4000)P_{yy}(2000, 4000) - [P_{xy}(2000, 4000)]^2 = (-0.002)^2 - (-0.001)^2 > 0$$

you can conclude that the production level of $x = 2000$ units and $y = 4000$ units yields a *maximum* profit. The maximum profit is

$$\begin{aligned} P(2000, 4000) &= 8(2000) + 10(4000) - (0.001)[2000^2 + 2000(4000) + 4000^2] - 10{,}000 \\ &= \$18{,}000. \end{aligned}$$

Try It Exploration A

NOTE In Example 2, it was assumed that the manufacturing plant is able to produce the required number of units to yield a maximum profit. In actual practice, the production would be bounded by physical constraints. You will study such constrained optimization problems in the next section.

The Method of Least Squares

Many of the examples in this text have involved **mathematical models.** For instance, Example 2 involves a quadratic model for profit. There are several ways to develop such models; one is called the **method of least squares.**

In constructing a model to represent a particular phenomenon, the goals are simplicity and accuracy. Of course, these goals often conflict. For instance, a simple linear model for the points in Figure 13.73 is

$$y = 1.8566x - 5.0246.$$

However, Figure 13.74 shows that by choosing the slightly more complicated quadratic model*

$$y = 0.1996x^2 - 0.7281x + 1.3749$$

you can achieve greater accuracy.

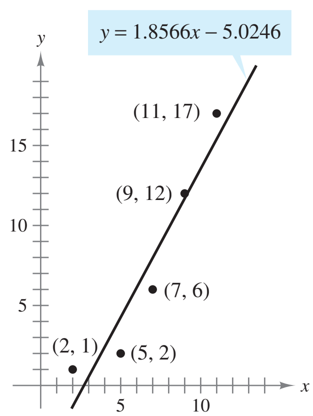


Figure 13.73

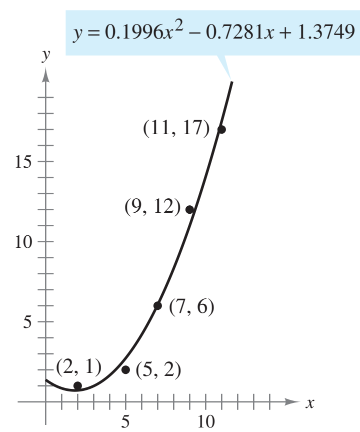


Figure 13.74

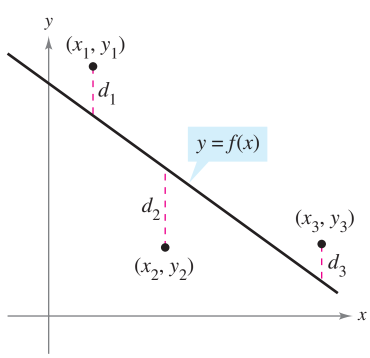
As a measure of how well the model $y = f(x)$ fits the collection of points

$$\{(x_1, y_1), (x_2, y_2), (x_3, y_3), \ldots, (x_n, y_n)\}$$

you can add the squares of the differences between the actual y-values and the values given by the model to obtain the **sum of the squared errors**

$$S = \sum_{i=1}^{n} [f(x_i) - y_i]^2.$$

Sum of the squared errors



Sum of the squared errors:
$S = d_1^2 + d_2^2 + d_3^2$
Figure 13.75

Graphically, S can be interpreted as the sum of the squares of the vertical distances between the graph of f and the given points in the plane, as shown in Figure 13.75. If the model is perfect, then $S = 0$. However, when perfection is not feasible, you can settle for a model that minimizes S. For instance, the sum of the squared errors for the linear model in Figure 13.73 is $S \approx 17$. Statisticians call the *linear model* that minimizes S the **least squares regression line.** The proof that this line actually minimizes S involves the minimizing of a function of two variables.

---

* *A method for finding the least squares quadratic model for a collection of data is described in Exercise 39.*

ADRIEN-MARIE LEGENDRE (1752–1833)

The method of least squares was introduced by the French mathematician Adrien-Marie Legendre. Legendre is best known for his work in geometry. In fact, his text Elements of Geometry was so popular in the United States that it continued to be used for 33 editions, spanning a period of more than 100 years.

MathBio

THEOREM 13.18 Least Squares Regression Line

The **least squares regression line** for $\{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}$ is given by $f(x) = ax + b$, where

$$a = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2} \quad \text{and} \quad b = \frac{1}{n}\left(\sum_{i=1}^{n} y_i - a\sum_{i=1}^{n} x_i\right).$$

Proof Let $S(a, b)$ represent the sum of the squared errors for the model $f(x) = ax + b$ and the given set of points. That is,

$$\begin{aligned} S(a, b) &= \sum_{i=1}^{n} [f(x_i) - y_i]^2 \\ &= \sum_{i=1}^{n} (ax_i + b - y_i)^2 \end{aligned}$$

where the points (x_i, y_i) represent constants. Because S is a function of a and b, you can use the methods discussed in the preceding section to find the minimum value of S. Specifically, the first partial derivatives of S are

$$\begin{aligned} S_a(a, b) &= \sum_{i=1}^{n} 2x_i(ax_i + b - y_i) \\ &= 2a\sum_{i=1}^{n} x_i^2 + 2b\sum_{i=1}^{n} x_i - 2\sum_{i=1}^{n} x_i y_i \\ S_b(a, b) &= \sum_{i=1}^{n} 2(ax_i + b - y_i) \\ &= 2a\sum_{i=1}^{n} x_i + 2nb - 2\sum_{i=1}^{n} y_i. \end{aligned}$$

By setting these two partial derivatives equal to 0, you obtain the values for a and b that are listed in the theorem. It is left to you to apply the Second Partials Test (see Exercise 40) to verify that these values of a and b yield a minimum.

If the x-values are symmetrically spaced about the y-axis, then $\Sigma\, x_i = 0$ and the formulas for a and b simplify to

$$a = \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2}$$

and

$$b = \frac{1}{n}\sum_{i=1}^{n} y_i.$$

This simplification is often possible with a translation of the x-values. For instance, if the x-values in a data collection consist of the years 2003, 2004, 2005, 2006, and 2007, you could let 2005 be represented by 0.

EXAMPLE 3 Finding the Least Squares Regression Line

Find the least squares regression line for the points $(-3, 0)$, $(-1, 1)$, $(0, 2)$, and $(2, 3)$.

Solution The table shows the calculations involved in finding the least squares regression line using $n = 4$.

x	y	xy	x^2
-3	0	0	9
-1	1	-1	1
0	2	0	0
2	3	6	4
$\sum_{i=1}^{n} x_i = -2$	$\sum_{i=1}^{n} y_i = 6$	$\sum_{i=1}^{n} x_i y_i = 5$	$\sum_{i=1}^{n} x_i^2 = 14$

TECHNOLOGY Many calculators have "built-in" least squares regression programs. If your calculator has such a program, use it to duplicate the results of Example 3.

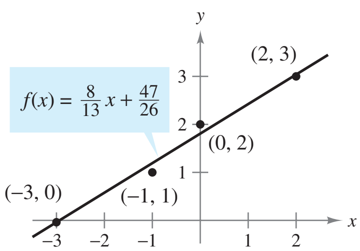
Applying Theorem 13.18 produces

$$a = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2} = \frac{4(5) - (-2)(6)}{4(14) - (-2)^2} = \frac{8}{13}$$

and

$$b = \frac{1}{n}\left(\sum_{i=1}^{n} y_i - a\sum_{i=1}^{n} x_i\right) = \frac{1}{4}\left[6 - \frac{8}{13}(-2)\right] = \frac{47}{26}.$$

The least squares regression line is $f(x) = \frac{8}{13}x + \frac{47}{26}$, as shown in Figure 13.76.



Least squares regression line
Figure 13.76

Editable Graph

Try It | Exploration A | Exploration B | Exploration C

Section 13.10 Lagrange Multipliers

- **Understand the Method of Lagrange Multipliers.**
- **Use Lagrange multipliers to solve constrained optimization problems.**
- **Use the Method of Lagrange Multipliers with two constraints.**

Lagrange Multipliers

Many optimization problems have restrictions, or **constraints,** on the values that can be used to produce the optimal solution. Such constraints tend to complicate optimization problems because the optimal solution can occur at a boundary point of the domain. In this section, you will study an ingenious technique for solving such problems. It is called the **Method of Lagrange Multipliers.**

To see how this technique works, suppose you want to find the rectangle of maximum area that can be inscribed in the ellipse given by

$$\frac{x^2}{3^2} + \frac{y^2}{4^2} = 1.$$

Let (x, y) be the vertex of the rectangle in the first quadrant, as shown in Figure 13.77. Because the rectangle has sides of lengths $2x$ and $2y$, its area is given by

$$f(x, y) = 4xy. \qquad \text{Objective function}$$

You want to find x and y such that $f(x, y)$ is a maximum. Your choice of (x, y) is restricted to first-quadrant points that lie on the ellipse

$$\frac{x^2}{3^2} + \frac{y^2}{4^2} = 1. \qquad \text{Constraint}$$

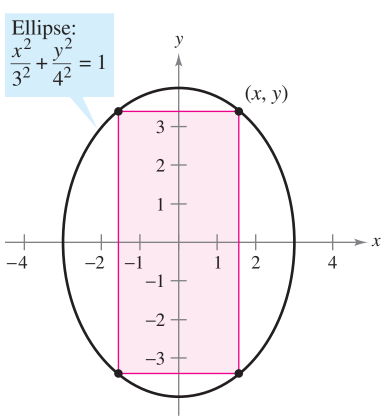
Now, consider the constraint equation to be a fixed level curve of

$$g(x, y) = \frac{x^2}{3^2} + \frac{y^2}{4^2}.$$

The level curves of f represent a family of hyperbolas

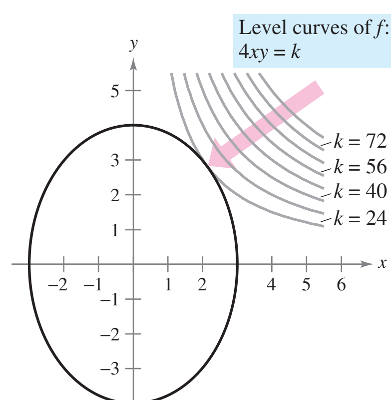
$$f(x, y) = 4xy = k.$$

In this family, the level curves that meet the given constraint correspond to the hyperbolas that intersect the ellipse. Moreover, to maximize $f(x, y)$, you want to find the hyperbola that just barely satisfies the constraint. The level curve that does this is the one that is *tangent* to the ellipse, as shown in Figure 13.78.



Objective function: $f(x, y) = 4xy$

Figure 13.77



Constraint: $g(x, y) = \frac{x^2}{3^2} + \frac{y^2}{4^2} = 1$

Figure 13.78

To find the appropriate hyperbola, use the fact that two curves are tangent at a point if and only if their gradient vectors are parallel. This means that $\nabla f(x, y)$ must be a scalar multiple of $\nabla g(x, y)$ at the point of tangency. In the context of constrained optimization problems, this scalar is denoted by λ (the lowercase Greek letter lambda).

$$\nabla f(x, y) = \lambda \nabla g(x, y)$$

The scalar λ is called a **Lagrange multiplier.** Theorem 13.19 gives the necessary conditions for the existence of such multipliers.

THEOREM 13.19 Lagrange's Theorem

Let f and g have continuous first partial derivatives such that f has an extremum at a point (x_0, y_0) on the smooth constraint curve $g(x, y) = c$. If $\nabla g(x_0, y_0) \neq \mathbf{0}$, then there is a real number λ such that

$$\nabla f(x_0, y_0) = \lambda \nabla g(x_0, y_0).$$

Joseph-Louis Lagrange (1736–1813)

The Method of Lagrange Multipliers is named after the French mathematician Joseph-Louis Lagrange. Lagrange first introduced the method in his famous paper on mechanics, written when he was just 19 years old.

Proof To begin, represent the smooth curve given by $g(x, y) = c$ by the vector-valued function

$$\mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j}, \quad \mathbf{r}'(t) \neq \mathbf{0}$$

where x' and y' are continuous on an open interval I. Define the function h as $h(t) = f(x(t), y(t))$. Then, because $f(x_0, y_0)$ is an extreme value of f, you know that

$$h(t_0) = f(x(t_0), y(t_0)) = f(x_0, y_0)$$

is an extreme value of h. This implies that $h'(t_0) = 0$, and, by the Chain Rule,

$$h'(t_0) = f_x(x_0, y_0)x'(t_0) + f_y(x_0, y_0)y'(t_0) = \nabla f(x_0, y_0) \cdot \mathbf{r}'(t_0) = 0.$$

So, $\nabla f(x_0, y_0)$ is orthogonal to $\mathbf{r}'(t_0)$. Moreover, by Theorem 13.12, $\nabla g(x_0, y_0)$ is also orthogonal to $\mathbf{r}'(t_0)$. Consequently, the gradients $\nabla f(x_0, y_0)$ and $\nabla g(x_0, y_0)$ are parallel, and there must exist a scalar λ such that

$$\nabla f(x_0, y_0) = \lambda \nabla g(x_0, y_0).$$

NOTE Lagrange's Theorem can be shown to be true for functions of three variables, using a similar argument with level surfaces and Theorem 13.14.

The Method of Lagrange Multipliers uses Theorem 13.19 to find the extreme values of a function f subject to a constraint.

Method of Lagrange Multipliers

Let f and g satisfy the hypothesis of Lagrange's Theorem, and let f have a minimum or maximum subject to the constraint $g(x, y) = c$. To find the minimum or maximum of f, use the following steps.

1. Simultaneously solve the equations $\nabla f(x, y) = \lambda \nabla g(x, y)$ and $g(x, y) = c$ by solving the following system of equations.

$$f_x(x, y) = \lambda g_x(x, y)$$
$$f_y(x, y) = \lambda g_y(x, y)$$
$$g(x, y) = c$$

2. Evaluate f at each solution point obtained in the first step. The largest value yields the maximum of f subject to the constraint $g(x, y) = c$, and the smallest value yields the minimum of f subject to the constraint $g(x, y) = c$.

NOTE As you will see in Examples 1 and 2, the Method of Lagrange Multipliers requires solving systems of nonlinear equations. This often can require some tricky algebraic manipulation.

Constrained Optimization Problems

In the problem at the beginning of this section, you wanted to maximize the area of a rectangle that is inscribed in an ellipse. Example 1 shows how to use Lagrange multipliers to solve this problem.

EXAMPLE 1 Using a Lagrange Multiplier with One Constraint

Find the maximum value of $f(x, y) = 4xy$ where $x > 0$ and $y > 0$, subject to the constraint $(x^2/3^2) + (y^2/4^2) = 1$.

NOTE Example 1 can also be solved using the techniques you learned in Chapter 3. To see how, try to find the maximum value of $A = 4xy$ given that

$$\frac{x^2}{3^2} + \frac{y^2}{4^2} = 1.$$

To begin, solve the second equation for y to obtain

$$y = \tfrac{4}{3}\sqrt{9 - x^2}.$$

Then substitute into the first equation to obtain

$$A = 4x\left(\tfrac{4}{3}\sqrt{9 - x^2}\right).$$

Finally, use the techniques of Chapter 3 to maximize A.

Solution To begin, let

$$g(x, y) = \frac{x^2}{3^2} + \frac{y^2}{4^2} = 1.$$

By equating $\nabla f(x, y) = 4y\mathbf{i} + 4x\mathbf{j}$ and $\lambda\nabla g(x, y) = (2\lambda x/9)\mathbf{i} + (\lambda y/8)\mathbf{j}$, you can obtain the following system of equations.

$$4y = \frac{2}{9}\lambda x \qquad f_x(x, y) = \lambda g_x(x, y)$$

$$4x = \frac{1}{8}\lambda y \qquad f_y(x, y) = \lambda g_y(x, y)$$

$$\frac{x^2}{3^2} + \frac{y^2}{4^2} = 1 \qquad \text{Constraint}$$

From the first equation, you obtain $\lambda = 18y/x$, and substitution into the second equation produces

$$4x = \frac{1}{8}\left(\frac{18y}{x}\right)y \quad \Rightarrow \quad x^2 = \frac{9}{16}y^2.$$

Substituting this value for x^2 into the third equation produces

$$\frac{1}{9}\left(\frac{9}{16}y^2\right) + \frac{1}{16}y^2 = 1 \quad \Rightarrow \quad y^2 = 8.$$

So, $y = \pm 2\sqrt{2}$. Because it is required that $y > 0$, choose the positive value and find that

$$\begin{aligned} x^2 &= \frac{9}{16}y^2 \\ &= \frac{9}{16}(8) = \frac{9}{2} \\ x &= \frac{3}{\sqrt{2}}. \end{aligned}$$

So, the maximum value of f is

$$f\left(\frac{3}{\sqrt{2}}, 2\sqrt{2}\right) = 4xy = 4\left(\frac{3}{\sqrt{2}}\right)\left(2\sqrt{2}\right) = 24.$$

Try It Exploration A Exploration B

Note that writing the constraint as

$$g(x, y) = \frac{x^2}{3^2} + \frac{y^2}{4^2} = 1 \qquad \text{or} \qquad g(x, y) = \frac{x^2}{3^2} + \frac{y^2}{4^2} - 1 = 0$$

does not affect the solution—the constant is eliminated when you form ∇g.

EXAMPLE 2 A Business Application

The Cobb-Douglas production function (see Example 5, Section 13.1) for a software manufacturer is given by

$$f(x, y) = 100x^{3/4}y^{1/4} \qquad \text{Objective function}$$

where x represents the units of labor (at \$150 per unit) and y represents the units of capital (at \$250 per unit). The total cost of labor and capital is limited to \$50,000. Find the maximum production level for this manufacturer.

FOR FURTHER INFORMATION For more information on the use of Lagrange multipliers in economics, see the article "Lagrange Multiplier Problems in Economics" by John V. Baxley and John C. Moorhouse in *The American Mathematical Monthly*.

MathArticle

Solution From the given function, you have

$$\nabla f(x, y) = 75x^{-1/4}y^{1/4}\,\mathbf{i} + 25x^{3/4}y^{-3/4}\,\mathbf{j}.$$

The limit on the cost of labor and capital produces the constraint

$$g(x, y) = 150x + 250y = 50{,}000. \qquad \text{Constraint}$$

So, $\lambda\nabla g(x, y) = 150\lambda\,\mathbf{i} + 250\lambda\,\mathbf{j}$. This gives rise to the following system of equations.

$$\begin{aligned} 75x^{-1/4}y^{1/4} &= 150\lambda && f_x(x, y) = \lambda g_x(x, y)\\ 25x^{3/4}y^{-3/4} &= 250\lambda && f_y(x, y) = \lambda g_y(x, y)\\ 150x + 250y &= 50{,}000 && \text{Constraint} \end{aligned}$$

By solving for λ in the first equation

$$\lambda = \frac{75x^{-1/4}y^{1/4}}{150} = \frac{x^{-1/4}y^{1/4}}{2}$$

and substituting into the second equation, you obtain

$$\begin{aligned} 25x^{3/4}y^{-3/4} &= 250\left(\frac{x^{-1/4}y^{1/4}}{2}\right)\\ 25x &= 125y. && \text{Multiply by } x^{1/4}y^{3/4}. \end{aligned}$$

So, $x = 5y$. By substituting into the third equation, you have

$$\begin{aligned} 150(5y) + 250y &= 50{,}000\\ 1000y &= 50{,}000\\ y &= 50 \text{ units of capital}\\ x &= 250 \text{ units of labor.} \end{aligned}$$

So, the maximum production level is

$$\begin{aligned} f(250, 50) &= 100(250)^{3/4}(50)^{1/4}\\ &\approx 16{,}719 \text{ product units.} \end{aligned}$$

Try It Exploration A

Economists call the Lagrange multiplier obtained in a production function the **marginal productivity of money.** For instance, in Example 2 the marginal productivity of money at $x = 250$ and $y = 50$ is

$$\lambda = \frac{x^{-1/4}y^{1/4}}{2} = \frac{(250)^{-1/4}(50)^{1/4}}{2} \approx 0.334$$

which means that for each additional dollar spent on production, an additional 0.334 unit of the product can be produced.

EXAMPLE 3 Lagrange Multipliers and Three Variables

Find the minimum value of

$$f(x, y, z) = 2x^2 + y^2 + 3z^2 \qquad \text{Objective function}$$

subject to the constraint $2x - 3y - 4z = 49$.

Solution Let $g(x, y, z) = 2x - 3y - 4z = 49$. Then, because

$$\nabla f(x, y, z) = 4x\mathbf{i} + 2y\mathbf{j} + 6z\mathbf{k} \quad \text{and} \quad \lambda\nabla g(x, y, z) = 2\lambda\mathbf{i} - 3\lambda\mathbf{j} - 4\lambda\mathbf{k}$$

you obtain the following system of equations.

$$\begin{aligned} 4x &= 2\lambda & & f_x(x, y, z) = \lambda g_x(x, y, z) \\ 2y &= -3\lambda & & f_y(x, y, z) = \lambda g_y(x, y, z) \\ 6z &= -4\lambda & & f_z(x, y, z) = \lambda g_z(x, y, z) \\ 2x - 3y - 4z &= 49 & & \text{Constraint} \end{aligned}$$

The solution of this system is $x = 3$, $y = -9$, and $z = -4$. So, the optimum value of f is

$$\begin{aligned} f(3, -9, -4) &= 2(3)^2 + (-9)^2 + 3(-4)^2 \\ &= 147. \end{aligned}$$

From the original function and constraint, it is clear that $f(x, y, z)$ has no maximum. So, the optimum value of f determined above is a minimum.

Try It **Exploration A** **Open Exploration**

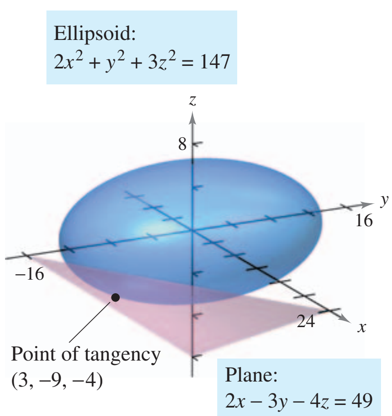


Figure 13.79

Rotatable Graph

A graphical interpretation of constrained optimization problems in two variables was given at the beginning of this section. In three variables, the interpretation is similar, except that level surfaces are used instead of level curves. For instance, in Example 3, the level surfaces of f are ellipsoids centered at the origin, and the constraint

$$2x - 3y - 4z = 49$$

is a plane. The minimum value of f is represented by the ellipsoid that is tangent to the constraint plane, as shown in Figure 13.79.

EXAMPLE 4 Optimization Inside a Region

Find the extreme values of

$$f(x, y) = x^2 + 2y^2 - 2x + 3 \qquad \text{Objective function}$$

subject to the constraint $x^2 + y^2 \le 10$.

Solution To solve this problem, you can break the constraint into two cases.

a. For points *on the circle* $x^2 + y^2 = 10$, you can use Lagrange multipliers to find that the maximum value of $f(x, y)$ is 24—this value occurs at $(-1, 3)$ and at $(-1, -3)$. In a similar way, you can determine that the minimum value of $f(x, y)$ is approximately 6.675—this value occurs at $(\sqrt{10}, 0)$.

b. For points *inside the circle*, you can use the techniques discussed in Section 13.8 to conclude that the function has a relative minimum of 2 at the point $(1, 0)$.

By combining these two results, you can conclude that f has a maximum of 24 at $(-1, \pm 3)$ and a minimum of 2 at $(1, 0)$, as shown in Figure 13.80.

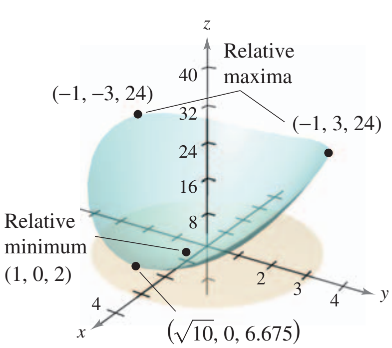


Figure 13.80

Rotatable Graph

Try It **Exploration A**

The Method of Lagrange Multipliers with Two Constraints

For optimization problems involving *two* constraint functions g and h, you can introduce a second Lagrange multiplier, μ (the lowercase Greek letter mu), and then solve the equation

$$\nabla f = \lambda \nabla g + \mu \nabla h$$

where the gradient vectors are not parallel, as illustrated in Example 5.

EXAMPLE 5 Optimization with Two Constraints

Let $T(x, y, z) = 20 + 2x + 2y + z^2$ represent the temperature at each point on the sphere $x^2 + y^2 + z^2 = 11$. Find the extreme temperatures on the curve formed by the intersection of the plane $x + y + z = 3$ and the sphere.

Solution The two constraints are

$$g(x, y, z) = x^2 + y^2 + z^2 = 11 \qquad \text{and} \qquad h(x, y, z) = x + y + z = 3.$$

Using

$$\begin{aligned} \nabla T(x, y, z) &= 2\mathbf{i} + 2\mathbf{j} + 2z\mathbf{k} \\ \lambda \nabla g(x, y, z) &= 2\lambda x\mathbf{i} + 2\lambda y\mathbf{j} + 2\lambda z\mathbf{k} \end{aligned}$$

and

$$\mu \nabla h(x, y, z) = \mu\mathbf{i} + \mu\mathbf{j} + \mu\mathbf{k}$$

you can write the following system of equations.

$$\begin{aligned} 2 &= 2\lambda x + \mu && T_x(x, y, z) = \lambda g_x(x, y, z) + \mu h_x(x, y, z) \\ 2 &= 2\lambda y + \mu && T_y(x, y, z) = \lambda g_y(x, y, z) + \mu h_y(x, y, z) \\ 2z &= 2\lambda z + \mu && T_z(x, y, z) = \lambda g_z(x, y, z) + \mu h_z(x, y, z) \\ x^2 + y^2 + z^2 &= 11 && \text{Constraint 1} \\ x + y + z &= 3 && \text{Constraint 2} \end{aligned}$$

By subtracting the second equation from the first, you can obtain the following system.

$$\begin{aligned} \lambda(x - y) &= 0 \\ 2z(1 - \lambda) - \mu &= 0 \\ x^2 + y^2 + z^2 &= 11 \\ x + y + z &= 3 \end{aligned}$$

STUDY TIP The system of equations that arises in the Method of Lagrange Multipliers is not, in general, a linear system, and finding the solution often requires ingenuity.

From the first equation, you can conclude that $\lambda = 0$ or $x = y$. If $\lambda = 0$, you can show that the critical points are $(3, -1, 1)$ and $(-1, 3, 1)$. (Try doing this—it takes a little work.) If $\lambda \neq 0$, then $x = y$ and you can show that the critical points occur when $x = y = \left(3 \pm 2\sqrt{3}\right)/3$ and $z = \left(3 \mp 4\sqrt{3}\right)/3$. Finally, to find the optimal solutions, compare the temperatures at the four critical points.

$$\begin{aligned} T(3, -1, 1) = T(-1, 3, 1) &= 25 \\ T\left(\frac{3 - 2\sqrt{3}}{3}, \frac{3 - 2\sqrt{3}}{3}, \frac{3 + 4\sqrt{3}}{3}\right) &= \frac{91}{3} \approx 30.33 \\ T\left(\frac{3 + 2\sqrt{3}}{3}, \frac{3 + 2\sqrt{3}}{3}, \frac{3 - 4\sqrt{3}}{3}\right) &= \frac{91}{3} \approx 30.33 \end{aligned}$$

So, $T = 25$ is the minimum temperature and $T = \frac{91}{3}$ is the maximum temperature on the curve.

Try It **Exploration A**

Section 14.1 Iterated Integrals and Area in the Plane

- **Evaluate an iterated integral.**
- **Use an iterated integral to find the area of a plane region.**

Iterated Integrals

NOTE In Chapters 14 and 15, you will study several applications of integration involving functions of several variables. Chapter 14 is much like Chapter 7 in that it surveys the use of integration to find plane areas, volumes, surface areas, moments, and centers of mass.

In Chapter 13, you saw that it is meaningful to differentiate functions of several variables with respect to one variable while holding the other variables constant. You can *integrate* functions of several variables by a similar procedure. For example, if you are given the partial derivative

$$f_x(x, y) = 2xy$$

then, by considering y constant, you can integrate with respect to x to obtain

$$\begin{aligned}
f(x, y) &= \int f_x(x, y)\,dx && \text{Integrate with respect to } x.\\
&= \int 2xy\,dx && \text{Hold } y \text{ constant.}\\
&= y\int 2x\,dx && \text{Factor out constant } y.\\
&= y(x^2) + C(y) && \text{Antiderivative of } 2x \text{ is } x^2.\\
&= x^2y + C(y). && C(y) \text{ is a function of } y.
\end{aligned}$$

The "constant" of integration, $C(y)$, is a function of y. In other words, by integrating with respect to x, you are able to recover $f(x, y)$ only partially. The total recovery of a function of x and y from its partial derivatives is a topic you will study in Chapter 15. For now, we are more concerned with extending definite integrals to functions of several variables. For instance, by considering y constant, you can apply the Fundamental Theorem of Calculus to evaluate

$$\int_1^{2y} 2xy\,dx = x^2y\Big]_1^{2y} = (2y)^2y - (1)^2y = 4y^3 - y.$$

x is the variable of integration and y is fixed.

Replace x by the limits of integration.

The result is a function of y.

Similarly, you can integrate with respect to y by holding x fixed. Both procedures are summarized as follows.

$$\int_{h_1(y)}^{h_2(y)} f_x(x, y)\,dx = f(x, y)\Big]_{h_1(y)}^{h_2(y)} = f(h_2(y), y) - f(h_1(y), y) \qquad \text{With respect to } x$$

$$\int_{g_1(x)}^{g_2(x)} f_y(x, y)\,dy = f(x, y)\Big]_{g_1(x)}^{g_2(x)} = f(x, g_2(x)) - f(x, g_1(x)) \qquad \text{With respect to } y$$

Note that the variable of integration cannot appear in either limit of integration. For instance, it makes no sense to write

$$\int_0^x y\,dx.$$

EXAMPLE 1 Integrating with Respect to y

Evaluate $\displaystyle\int_1^x (2x^2y^{-2} + 2y)\,dy$.

Solution Considering x to be constant and integrating with respect to y produces

$$\int_1^x (2x^2y^{-2} + 2y)\,dy = \left[\frac{-2x^2}{y} + y^2\right]_1^x \qquad \text{Integrate with respect to } y.$$
$$= \left(\frac{-2x^2}{x} + x^2\right) - \left(\frac{-2x^2}{1} + 1\right)$$
$$= 3x^2 - 2x - 1.$$

Try It Exploration A

Notice in Example 1 that the integral defines a function of x and can *itself* be integrated, as shown in the next example.

EXAMPLE 2 The Integral of an Integral

Evaluate $\displaystyle\int_1^2 \left[\int_1^x (2x^2y^{-2} + 2y)\,dy\right] dx$.

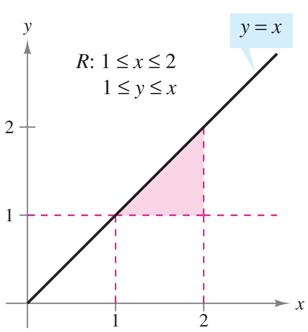
Solution Using the result of Example 1, you have

$$\int_1^2 \left[\int_1^x (2x^2y^{-2} + 2y)\,dy\right] dx = \int_1^2 (3x^2 - 2x - 1)\,dx$$
$$= \Big[x^3 - x^2 - x\Big]_1^2 \qquad \text{Integrate with respect to } x.$$
$$= 2 - (-1)$$
$$= 3.$$

Try It Exploration A

The integral in Example 2 is an **iterated integral.** The brackets used in Example 2 are normally not written. Instead, iterated integrals are usually written simply as

$$\int_a^b \int_{g_1(x)}^{g_2(x)} f(x, y)\,dy\,dx \quad \text{and} \quad \int_c^d \int_{h_1(y)}^{h_2(y)} f(x, y)\,dx\,dy.$$



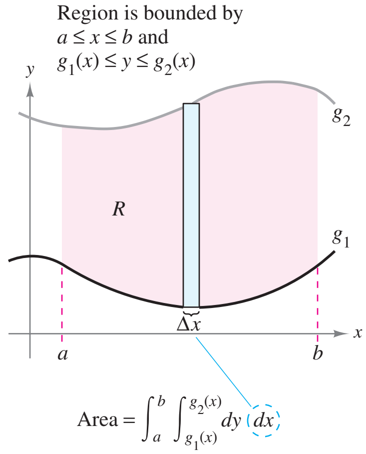
The region of integration for $\displaystyle\int_1^2 \int_1^x f(x, y)\,dy\,dx$

Figure 14.1

The **inside limits of integration** can be variable with respect to the outer variable of integration. However, the **outside limits of integration** *must be* constant with respect to both variables of integration. After performing the inside integration, you obtain a "standard" definite integral, and the second integration produces a real number. The limits of integration for an iterated integral identify two sets of boundary intervals for the variables. For instance, in Example 2, the outside limits indicate that x lies in the interval $1 \le x \le 2$ and the inside limits indicate that y lies in the interval $1 \le y \le x$. Together, these two intervals determine the **region of integration *R*** of the iterated integral, as shown in Figure 14.1.

Because an iterated integral is just a special type of definite integral—one in which the integrand is also an integral—you can use the properties of definite integrals to evaluate iterated integrals.

Area of a Plane Region



Vertically simple region
Figure 14.2

In the remainder of this section, you will take a new look at an old problem—that of finding the area of a plane region. Consider the plane region R bounded by $a \le x \le b$ and $g_1(x) \le y \le g_2(x)$, as shown in Figure 14.2. The area of R is given by the definite integral

$$\int_a^b [g_2(x) - g_1(x)]\, dx. \qquad \text{Area of } R$$

Using the Fundamental Theorem of Calculus, you can rewrite the integrand $g_2(x) - g_1(x)$ as a definite integral. Specifically, if you consider x to be fixed and let y vary from $g_1(x)$ to $g_2(x)$, you can write

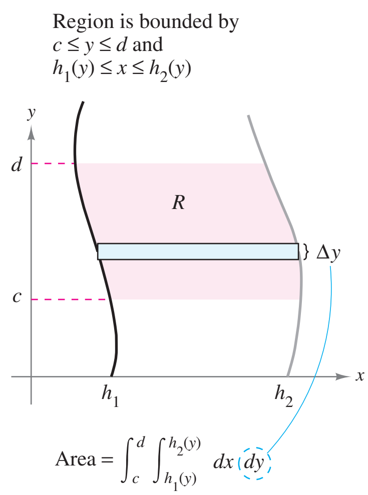
$$\int_{g_1(x)}^{g_2(x)} dy = y\Big]_{g_1(x)}^{g_2(x)} = g_2(x) - g_1(x).$$

Combining these two integrals, you can write the area of the region R as an iterated integral

$$\int_a^b \int_{g_1(x)}^{g_2(x)} dy\, dx = \int_a^b y\Big]_{g_1(x)}^{g_2(x)} dx \qquad \text{Area of } R$$
$$= \int_a^b [g_2(x) - g_1(x)]\, dx.$$

Placing a representative rectangle in the region R helps determine both the order and the limits of integration. A vertical rectangle implies the order $dy\, dx$, with the inside limits corresponding to the upper and lower bounds of the rectangle, as shown in Figure 14.2. This type of region is called **vertically simple,** because the outside limits of integration represent the vertical lines $x = a$ and $x = b$.

Similarly, a horizontal rectangle implies the order $dx\, dy$, with the inside limits determined by the left and right bounds of the rectangle, as shown in Figure 14.3. This type of region is called **horizontally simple,** because the outside limits represent the horizontal lines $y = c$ and $y = d$. The iterated integrals used for these two types of simple regions are summarized as follows.



Horizontally simple region
Figure 14.3

Area of a Region in the Plane

1. If R is defined by $a \le x \le b$ and $g_1(x) \le y \le g_2(x)$, where g_1 and g_2 are continuous on $[a, b]$, then the area of R is given by

$$A = \int_a^b \int_{g_1(x)}^{g_2(x)} dy\, dx. \qquad \text{Figure 14.2 (vertically simple)}$$

2. If R is defined by $c \le y \le d$ and $h_1(y) \le x \le h_2(y)$, where h_1 and h_2 are continuous on $[c, d]$, then the area of R is given by

$$A = \int_c^d \int_{h_1(y)}^{h_2(y)} dx\, dy. \qquad \text{Figure 14.3 (horizontally simple)}$$

NOTE Be sure you see that the order of integration of these two integrals is different—the order $dy\, dx$ corresponds to a vertically simple region, and the order $dx\, dy$ corresponds to a horizontally simple region.

If all four limits of integration happen to be constants, the region of integration is rectangular, as shown in Example 3.

EXAMPLE 3 The Area of a Rectangular Region

Use an iterated integral to represent the area of the rectangle shown in Figure 14.4.

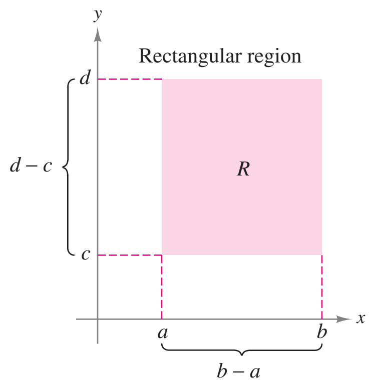


Figure 14.4

Solution The region shown in Figure 14.4 is both vertically simple and horizontally simple, so you can use either order of integration. By choosing the order $dy\,dx$, you obtain the following.

$$\begin{aligned}
\int_a^b \int_c^d dy\,dx &= \int_a^b y\Big]_c^d dx && \text{Integrate with respect to } y.\\
&= \int_a^b (d-c)\,dx\\
&= \Big[(d-c)x\Big]_a^b && \text{Integrate with respect to } x.\\
&= (d-c)(b-a)
\end{aligned}$$

Notice that this answer is consistent with what you know from geometry.

Exploration A Exploration B

EXAMPLE 4 Finding Area by an Iterated Integral

Use an iterated integral to find the area of the region bounded by the graphs of

$$f(x) = \sin x \qquad \text{Sine curve forms upper boundary.}$$
$$g(x) = \cos x \qquad \text{Cosine curve forms lower boundary.}$$

between $x = \pi/4$ and $x = 5\pi/4$.

Solution Because f and g are given as functions of x, a vertical representative rectangle is convenient, and you can choose $dy\,dx$ as the order of integration, as shown in Figure 14.5. The outside limits of integration are $\pi/4 \le x \le 5\pi/4$. Moreover, because the rectangle is bounded above by $f(x) = \sin x$ and below by $g(x) = \cos x$, you have

$$\begin{aligned}
\text{Area of } R &= \int_{\pi/4}^{5\pi/4} \int_{\cos x}^{\sin x} dy\,dx\\
&= \int_{\pi/4}^{5\pi/4} y\Big]_{\cos x}^{\sin x} dx && \text{Integrate with respect to } y.\\
&= \int_{\pi/4}^{5\pi/4} (\sin x - \cos x)\,dx\\
&= \Big[-\cos x - \sin x\Big]_{\pi/4}^{5\pi/4} && \text{Integrate with respect to } x.\\
&= 2\sqrt{2}.
\end{aligned}$$

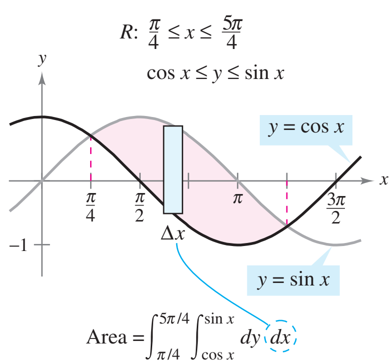
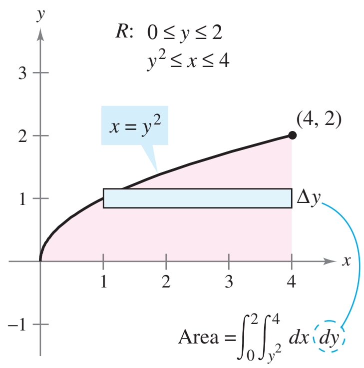


Figure 14.5

Editable Graph

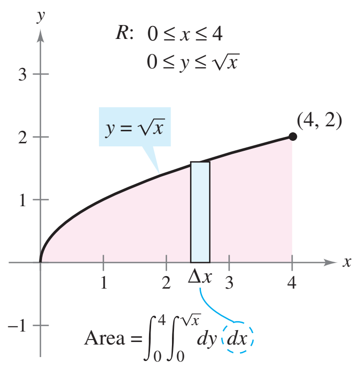
Try It Exploration A Exploration B

NOTE The region of integration of an iterated integral need not have any straight lines as boundaries. For instance, the region of integration shown in Figure 14.5 is *vertically simple* even though it has no vertical lines as left and right boundaries. The quality that makes the region vertically simple is that it is bounded above and below by the graphs of *functions of x*.



(a)

Editable Graph



(b)
Figure 14.6

One order of integration will often produce a simpler integration problem than the other order. For instance, try reworking Example 4 with the order *dx dy*—you may be surprised to see that the task is formidable. However, if you succeed, you will see that the answer is the same. In other words, the order of integration affects the ease of integration, but not the value of the integral.

EXAMPLE 5 Comparing Different Orders of Integration

Sketch the region whose area is represented by the integral

$$\int_0^2 \int_{y^2}^4 dx\, dy.$$

Then find another iterated integral using the order *dy dx* to represent the same area and show that both integrals yield the same value.

Solution From the given limits of integration, you know that

$$y^2 \le x \le 4 \qquad \text{Inner limits of integration}$$

which means that the region R is bounded on the left by the parabola $x = y^2$ and on the right by the line $x = 4$. Furthermore, because

$$0 \le y \le 2 \qquad \text{Outer limits of integration}$$

you know that R is bounded below by the x-axis, as shown in Figure 14.6(a). The value of this integral is

$$\int_0^2 \int_{y^2}^4 dx\, dy = \int_0^2 x \Big]_{y^2}^4 dy \qquad \text{Integrate with respect to } x.$$

$$= \int_0^2 (4 - y^2)\, dy$$

$$= \left[4y - \frac{y^3}{3}\right]_0^2 = \frac{16}{3}. \qquad \text{Integrate with respect to } y.$$

To change the order of integration to *dy dx*, place a vertical rectangle in the region, as shown in Figure 14.6(b). From this you can see that the constant bounds $0 \le x \le 4$ serve as the outer limits of integration. By solving for y in the equation $x = y^2$, you can conclude that the inner bounds are $0 \le y \le \sqrt{x}$. So, the area of the region can also be represented by

$$\int_0^4 \int_0^{\sqrt{x}} dy\, dx.$$

By evaluating this integral, you can see that it has the same value as the original integral.

$$\int_0^4 \int_0^{\sqrt{x}} dy\, dx = \int_0^4 y \Big]_0^{\sqrt{x}} dx \qquad \text{Integrate with respect to } y.$$

$$= \int_0^4 \sqrt{x}\, dx$$

$$= \frac{2}{3} x^{3/2} \Big]_0^4 = \frac{16}{3} \qquad \text{Integrate with respect to } x.$$

Try It | Exploration A | Exploration B | Open Exploration

Sometimes it is not possible to calculate the area of a region with a single iterated integral. In these cases you can divide the region into subregions such that the area of each subregion can be calculated by an iterated integral. The total area is then the sum of the iterated integrals.

TECHNOLOGY Some computer software can perform symbolic integration for integrals such as those in Example 6. If you have access to such software, use it to evaluate the integrals in the exercises and examples given in this section.

EXAMPLE 6 An Area Represented by Two Iterated Integrals

Find the area of the region R that lies below the parabola

$$y = 4x - x^2 \qquad \text{Parabola forms upper boundary.}$$

above the x-axis, and above the line

$$y = -3x + 6. \qquad \text{Line and } x\text{-axis form lower boundary.}$$

Solution Begin by dividing R into the two subregions R_1 and R_2 shown in Figure 14.7.

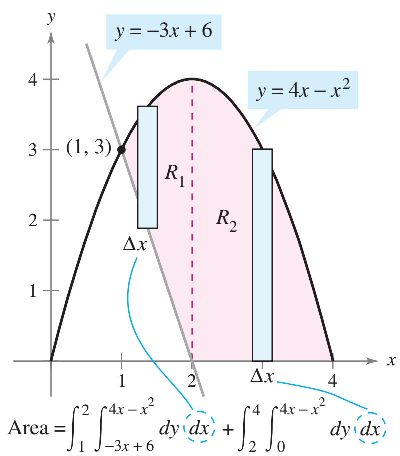


Figure 14.7

Editable Graph

In both regions, it is convenient to use vertical rectangles, and you have

$$\begin{aligned}
\text{Area} &= \int_1^2 \int_{-3x+6}^{4x-x^2} dy\,dx + \int_2^4 \int_0^{4x-x^2} dy\,dx \\
&= \int_1^2 (4x - x^2 + 3x - 6)\,dx + \int_2^4 (4x - x^2)\,dx \\
&= \left[\frac{7x^2}{2} - \frac{x^3}{3} - 6x\right]_1^2 + \left[2x^2 - \frac{x^3}{3}\right]_2^4 \\
&= \left(14 - \frac{8}{3} - 12 - \frac{7}{2} + \frac{1}{3} + 6\right) + \left(32 - \frac{64}{3} - 8 + \frac{8}{3}\right) = \frac{15}{2}.
\end{aligned}$$

The area of the region is $15/2$ square units. Try checking this using the procedure for finding the area between two curves, as presented in Section 7.1.

Try It Exploration A

At this point you may be wondering why you would need iterated integrals. After all, you already know how to use conventional integration to find the area of a region in the plane. (For instance, compare the solution of Example 4 in this section with that given in Example 3 in Section 7.1.) The need for iterated integrals will become clear in the next section. In this section, primary attention is given to procedures for finding the limits of integration of the region of an iterated integral, and the following exercise set is designed to develop skill in this important procedure.

NOTE In Examples 3 to 6, be sure you see the benefit of sketching the region of integration. You should develop the habit of making sketches to help determine the limits of integration for all iterated integrals in this chapter.

Section 14.2 Double Integrals and Volume

- **Use a double integral to represent the volume of a solid region.**
- **Use properties of double integrals.**
- **Evaluate a double integral as an iterated integral.**

Double Integrals and Volume of a Solid Region

You already know that a definite integral over an *interval* uses a limit process to assign measure to quantities such as area, volume, arc length, and mass. In this section, you will use a similar process to define the **double integral** of a function of two variables over a *region in the plane*.

Consider a continuous function f such that $f(x, y) \geq 0$ for all (x, y) in a region R in the xy-plane. The goal is to find the volume of the solid region lying between the surface given by

$$z = f(x, y) \qquad \text{Surface lying above the } xy\text{-plane}$$

and the xy-plane, as shown in Figure 14.8. You can begin by superimposing a rectangular grid over the region, as shown in Figure 14.9. The rectangles lying entirely within R form an **inner partition** Δ, whose **norm** $\|\Delta\|$ is defined as the length of the longest diagonal of the n rectangles. Next, choose a point (x_i, y_i) in each rectangle and form the rectangular prism whose height is $f(x_i, y_i)$, as shown in Figure 14.10. Because the area of the ith rectangle is

$$\Delta A_i \qquad \text{Area of } i\text{th rectangle}$$

it follows that the volume of the ith prism is

$$f(x_i, y_i)\,\Delta A_i \qquad \text{Volume of } i\text{th prism}$$

and you can approximate the volume of the solid region by the Riemann sum of the volumes of all n prisms,

$$\sum_{i=1}^{n} f(x_i, y_i)\,\Delta A_i \qquad \text{Riemann sum}$$

as shown in Figure 14.11. This approximation can be improved by tightening the mesh of the grid to form smaller and smaller rectangles, as shown in Example 1.

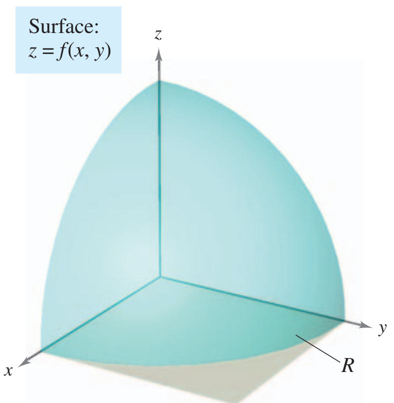
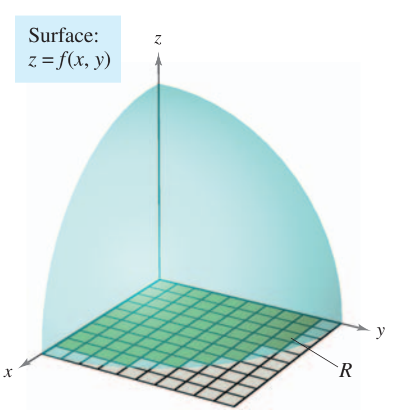


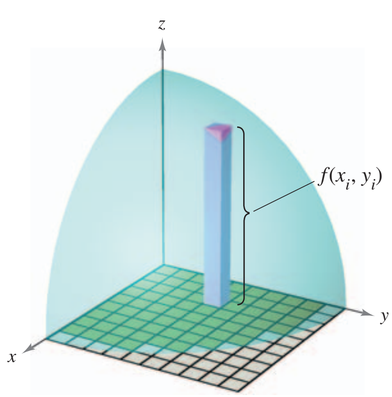
Figure 14.8

Rotatable Graph



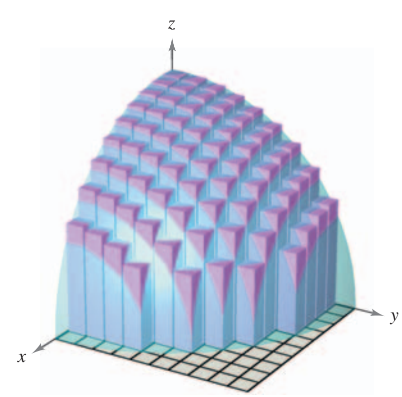
The rectangles lying within R form an inner partition of R.
Figure 14.9

Rotatable Graph



Rectangular prism whose base has an area of ΔA_i and whose height is $f(x_i, y_i)$
Figure 14.10

Rotatable Graph



Volume approximated by rectangular prisms
Figure 14.11

Rotatable Graph

EXAMPLE 1 Approximating the Volume of a Solid

Approximate the volume of the solid lying between the paraboloid

$$f(x, y) = 1 - \frac{1}{2}x^2 - \frac{1}{2}y^2$$

and the square region R given by $0 \le x \le 1$, $0 \le y \le 1$. Use a partition made up of squares whose sides have a length of $\frac{1}{4}$.

Solution Begin by forming the specified partition of R. For this partition, it is convenient to choose the centers of the subregions as the points at which to evaluate $f(x, y)$.

$$\begin{array}{cccc}
\left(\frac{1}{8}, \frac{1}{8}\right) & \left(\frac{1}{8}, \frac{3}{8}\right) & \left(\frac{1}{8}, \frac{5}{8}\right) & \left(\frac{1}{8}, \frac{7}{8}\right) \\
\left(\frac{3}{8}, \frac{1}{8}\right) & \left(\frac{3}{8}, \frac{3}{8}\right) & \left(\frac{3}{8}, \frac{5}{8}\right) & \left(\frac{3}{8}, \frac{7}{8}\right) \\
\left(\frac{5}{8}, \frac{1}{8}\right) & \left(\frac{5}{8}, \frac{3}{8}\right) & \left(\frac{5}{8}, \frac{5}{8}\right) & \left(\frac{5}{8}, \frac{7}{8}\right) \\
\left(\frac{7}{8}, \frac{1}{8}\right) & \left(\frac{7}{8}, \frac{3}{8}\right) & \left(\frac{7}{8}, \frac{5}{8}\right) & \left(\frac{7}{8}, \frac{7}{8}\right)
\end{array}$$

Because the area of each square is $\Delta A_i = \frac{1}{16}$, you can approximate the volume by the sum

$$\sum_{i=1}^{16} f(x_i\, y_i)\,\Delta A_i = \sum_{i=1}^{16}\left(1 - \frac{1}{2}x_i^2 - \frac{1}{2}y_i^2\right)\left(\frac{1}{16}\right) \approx 0.672.$$

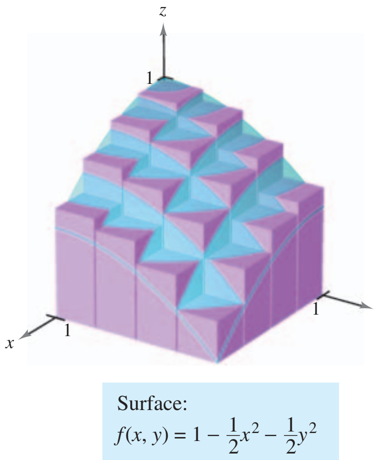


Figure 14.12

Rotatable Graph

This approximation is shown graphically in Figure 14.12. The exact volume of the solid is $\frac{2}{3}$ (see Example 2). You can obtain a better approximation by using a finer partition. For example, with a partition of squares with sides of length $\frac{1}{10}$, the approximation is 0.668.

Try It   Exploration A

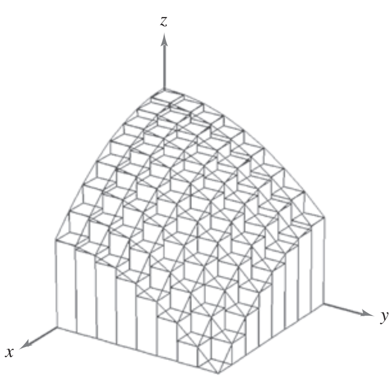


Figure 14.13

Rotatable Graph

TECHNOLOGY Some three-dimensional graphing utilities are capable of graphing figures such as that shown in Figure 14.12. For instance, the graph shown in Figure 14.13 was drawn with a computer program. In this graph, note that each of the rectangular prisms lies within the solid region.

In Example 1, note that by using finer partitions, you obtain better approximations of the volume. This observation suggests that you could obtain the exact volume by taking a limit. That is,

$$\text{Volume} = \lim_{\|\Delta\| \to 0} \sum_{i=1}^{n} f(x_i, y_i)\,\Delta A_i.$$

The precise meaning of this limit is that the limit is equal to L if for every $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$\left| L - \sum_{i=1}^{n} f(x_i, y_i)\,\Delta A_i \right| < \varepsilon$$

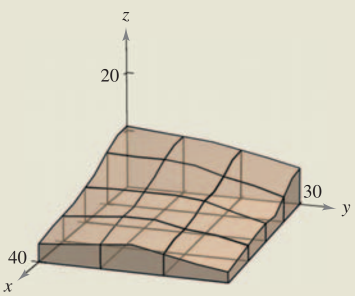
for all partitions Δ of the plane region R (that satisfy $\|\Delta\| < \delta$) and for all possible choices of x_i and y_i in the ith region.

Using the limit of a Riemann sum to define volume is a special case of using the limit to define a **double integral.** The general case, however, does not require that the function be positive or continuous.

EXPLORATION

The entries in the table represent the depth (in 10-yard units) of earth at the center of each square in the figure below.

x \ y	1	2	3
1	10	9	7
2	7	7	4
3	5	5	4
4	4	5	3



Rotatable Graph

Approximate the number of cubic yards of earth in the first octant. (This exploration was submitted by Robert Vojack, Ridgewood High School, Ridgewood, NJ.)

Definition of Double Integral

If f is defined on a closed, bounded region R in the xy-plane, then the **double integral of f over R** is given by

$$\int\!\!\int_R f(x, y)\, dA = \lim_{\|\Delta\| \to 0} \sum_{i=1}^{n} f(x_i, y_i)\, \Delta A_i$$

provided the limit exists. If the limit exists, then f is **integrable** over R.

NOTE Having defined a double integral, you will see that a definite integral is occasionally referred to as a **single integral.**

Sufficient conditions for the double integral of f on the region R to exist are that R can be written as a union of a finite number of nonoverlapping subregions (see Figure 14.14) that are vertically or horizontally simple *and* that f is continuous on the region R.

A double integral can be used to find the volume of a solid region that lies between the xy-plane and the surface given by $z = f(x, y)$.

Volume of a Solid Region

If f is integrable over a plane region R and $f(x, y) \ge 0$ for all (x, y) in R, then the volume of the solid region that lies above R and below the graph of f is defined as

$$V = \int\!\!\int_R f(x, y)\, dA.$$

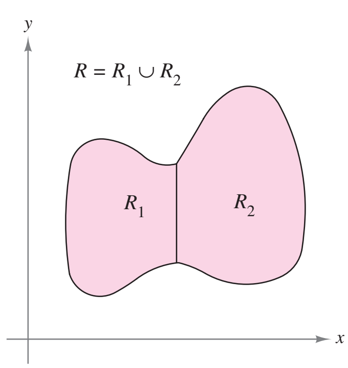
Properties of Double Integrals

Double integrals share many properties of single integrals.

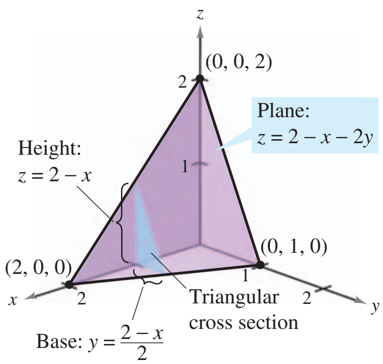
THEOREM 14.1 Properties of Double Integrals

Let f and g be continuous over a closed, bounded plane region R, and let c be a constant.

1. $\displaystyle\int\!\!\int_R cf(x, y)\, dA = c\int\!\!\int_R f(x, y)\, dA$
2. $\displaystyle\int\!\!\int_R [f(x, y) \pm g(x, y)]\, dA = \int\!\!\int_R f(x, y)\, dA \pm \int\!\!\int_R g(x, y)\, dA$
3. $\displaystyle\int\!\!\int_R f(x, y)\, dA \ge 0, \qquad \text{if } f(x, y) \ge 0$
4. $\displaystyle\int\!\!\int_R f(x, y)\, dA \ge \int\!\!\int_R g(x, y)\, dA, \qquad \text{if } f(x, y) \ge g(x, y)$
5. $\displaystyle\int\!\!\int_R f(x, y)\, dA = \int\!\!\int_{R_1} f(x, y)\, dA + \int\!\!\int_{R_2} f(x, y)\, dA$, where R is the union of two nonoverlapping subregions R_1 and R_2.



Two regions are nonoverlapping if their intersection is a set that has an area of 0. In this figure, the area of the line segment that is common to R_1 and R_2 is 0.
Figure 14.14



Volume: $\int_0^2 A(x)\,dx$

Figure 14.15

Rotatable Graph

Evaluation of Double Integrals

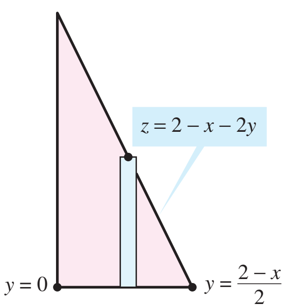
Normally, the first step in evaluating a double integral is to rewrite it as an iterated integral. To show how this is done, a geometric model of a double integral is used as the volume of a solid.

Consider the solid region bounded by the plane $z = f(x, y) = 2 - x - 2y$ and the three coordinate planes, as shown in Figure 14.15. Each vertical cross section taken parallel to the yz-plane is a triangular region whose base has a length of $y = (2 - x)/2$ and whose height is $z = 2 - x$. This implies that for a fixed value of x, the area of the triangular cross section is

$$A(x) = \frac{1}{2}(\text{base})(\text{height}) = \frac{1}{2}\left(\frac{2-x}{2}\right)(2-x) = \frac{(2-x)^2}{4}.$$

By the formula for the volume of a solid with known cross sections (Section 7.2), the volume of the solid is

$$\begin{aligned}\text{Volume} &= \int_a^b A(x)\,dx\\ &= \int_0^2 \frac{(2-x)^2}{4}\,dx\\ &= -\frac{(2-x)^3}{12}\bigg]_0^2 = \frac{2}{3}.\end{aligned}$$



Triangular cross section
Figure 14.16

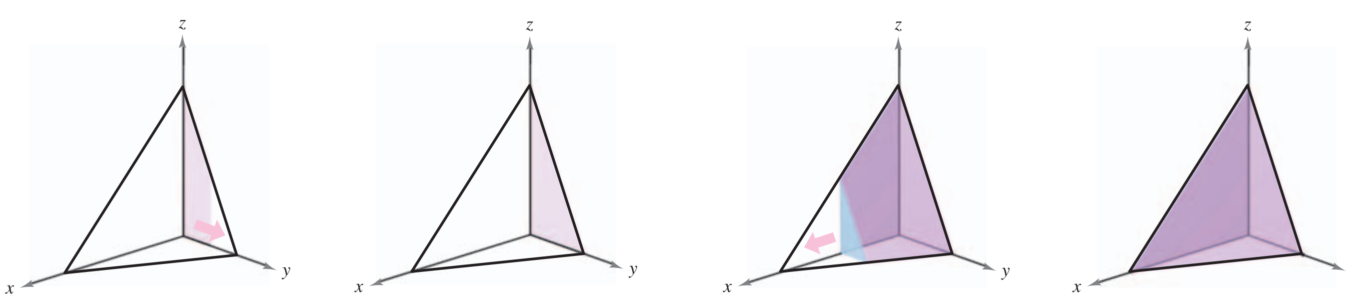
This procedure works no matter how $A(x)$ is obtained. In particular, you can find $A(x)$ by integration, as shown in Figure 14.16. That is, you consider x to be constant, and integrate $z = 2 - x - 2y$ from 0 to $(2 - x)/2$ to obtain

$$\begin{aligned}A(x) &= \int_0^{(2-x)/2} (2 - x - 2y)\,dy\\ &= \Big[(2-x)y - y^2\Big]_0^{(2-x)/2}\\ &= \frac{(2-x)^2}{4}.\end{aligned}$$

Combining these results, you have the *iterated integral*

$$\text{Volume} = \iint_R f(x, y)\,dA = \int_0^2 \int_0^{(2-x)/2} (2 - x - 2y)\,dy\,dx.$$

To understand this procedure better, it helps to imagine the integration as two sweeping motions. For the inner integration, a vertical line sweeps out the area of a cross section. For the outer integration, the triangular cross section sweeps out the volume, as shown in Figure 14.17.



Integrate with respect to y to obtain the area of the cross section. Integrate with respect to x to obtain the volume of the solid.
Figure 14.17

Animation

The following theorem was proved by the Italian mathematician Guido Fubini (1879–1943). The theorem states that if R is a vertically or horizontally simple region and f is continuous on R, the double integral of f on R is equal to an iterated integral.

THEOREM 14.2 Fubini's Theorem

Let f be continuous on a plane region R.

1. If R is defined by $a \le x \le b$ and $g_1(x) \le y \le g_2(x)$, where g_1 and g_2 are continuous on $[a, b]$, then

$$\int_R\int f(x, y)\, dA = \int_a^b \int_{g_1(x)}^{g_2(x)} f(x, y)\, dy\, dx.$$

2. If R is defined by $c \le y \le d$ and $h_1(y) \le x \le h_2(y)$, where h_1 and h_2 are continuous on $[c, d]$, then

$$\int_R\int f(x, y)\, dA = \int_c^d \int_{h_1(y)}^{h_2(y)} f(x, y)\, dx\, dy.$$

EXAMPLE 2 Evaluating a Double Integral as an Iterated Integral

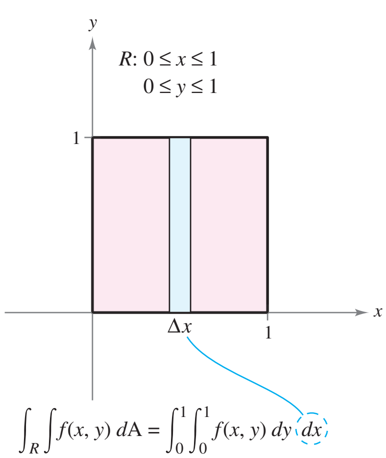
Evaluate

$$\int_R\int \left(1 - \frac{1}{2}x^2 - \frac{1}{2}y^2\right) dA$$

where R is the region given by $0 \le x \le 1$, $0 \le y \le 1$.

Solution Because the region R is a square, it is both vertically and horizontally simple, and you can use either order of integration. Choose $dy\, dx$ by placing a vertical representative rectangle in the region, as shown in Figure 14.18. This produces the following.

$$\begin{aligned}\int_R\int \left(1 - \frac{1}{2}x^2 - \frac{1}{2}y^2\right) dA &= \int_0^1\int_0^1 \left(1 - \frac{1}{2}x^2 - \frac{1}{2}y^2\right) dy\, dx\\ &= \int_0^1 \left[\left(1 - \frac{1}{2}x^2\right)y - \frac{y^3}{6}\right]_0^1 dx\\ &= \int_0^1 \left(\frac{5}{6} - \frac{1}{2}x^2\right) dx\\ &= \left[\frac{5}{6}x - \frac{x^3}{6}\right]_0^1\\ &= \frac{2}{3}\end{aligned}$$



The volume of the solid region is $\frac{2}{3}$.
Figure 14.18

Try It Exploration A

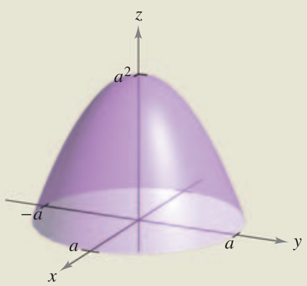
The double integral evaluated in Example 2 represents the volume of the solid region approximated in Example 1. Note that the approximation obtained in Example 1 is quite good $\left(0.672 \text{ vs. } \frac{2}{3}\right)$, even though you used a partition consisting of only 16 squares. The error resulted because the centers of the square subregions were used as the points in the approximation. This is comparable to the Midpoint Rule approximation of a single integral.

EXPLORATION

Volume of a Paraboloid Sector
The solid in Example 3 has an elliptical (not a circular) base. Consider the region bounded by the circular paraboloid

$$z = a^2 - x^2 - y^2, \quad a > 0$$

and the xy-plane. How many ways do you now know for finding the volume of this solid? For instance, you could use the disk method to find the volume as a solid of revolution. Does each method involve integration?



Rotatable Graph

The difficulty of evaluating a single integral $\int_a^b f(x)\,dx$ usually depends on the function f, and not on the interval $[a, b]$. This is a major difference between single and double integrals. In the next example, you will integrate a function similar to that in Examples 1 and 2. Notice that a change in the region R produces a much more difficult integration problem.

EXAMPLE 3 Finding Volume by a Double Integral

Find the volume of the solid region bounded by the paraboloid $z = 4 - x^2 - 2y^2$ and the xy-plane.

Solution By letting $z = 0$, you can see that the base of the region in the xy-plane is the ellipse $x^2 + 2y^2 = 4$, as shown in Figure 14.19(a). This plane region is both vertically and horizontally simple, so the order $dy\,dx$ is appropriate.

Variable bounds for y: $-\sqrt{\dfrac{(4 - x^2)}{2}} \le y \le \sqrt{\dfrac{(4 - x^2)}{2}}$

Constant bounds for x: $-2 \le x \le 2$

The volume is given by

$$V = \int_{-2}^{2}\int_{-\sqrt{(4-x^2)/2}}^{\sqrt{(4-x^2)/2}} (4 - x^2 - 2y^2)\,dy\,dx \qquad \text{See Figure 14.19(b).}$$

$$= \int_{-2}^{2}\left[(4 - x^2)y - \frac{2y^3}{3}\right]_{-\sqrt{(4-x^2)/2}}^{\sqrt{(4-x^2)/2}} dx$$

$$= \frac{4}{3\sqrt{2}}\int_{-2}^{2} (4 - x^2)^{3/2}\,dx$$

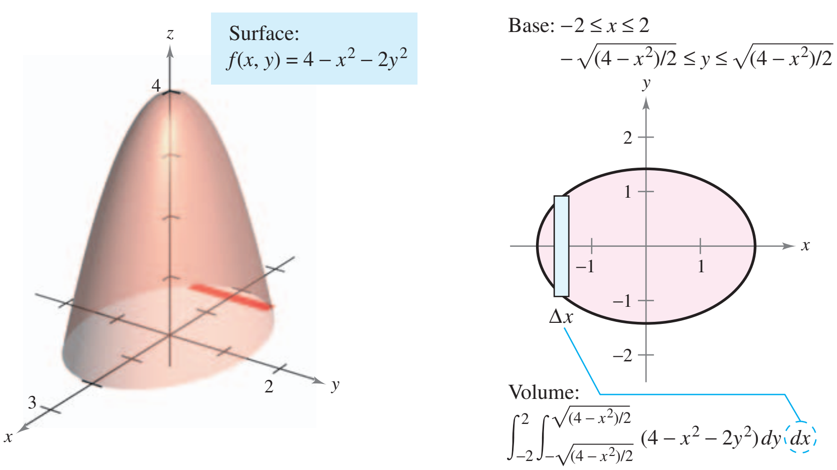
$$= \frac{4}{3\sqrt{2}}\int_{-\pi/2}^{\pi/2} 16\cos^4\theta\,d\theta \qquad x = 2\sin\theta$$

$$= \frac{64}{3\sqrt{2}}(2)\int_{0}^{\pi/2} \cos^4\theta\,d\theta$$

$$= \frac{128}{3\sqrt{2}}\left(\frac{3\pi}{16}\right) \qquad \text{Wallis's Formula}$$

$$= 4\sqrt{2}\pi.$$

NOTE In Example 3, note the usefulness of Wallis's Formula to evaluate $\int_0^{\pi/2} \cos^n\theta\,d\theta$. You may want to review this formula in Section 8.3.



(a) (b)

Rotatable Graph

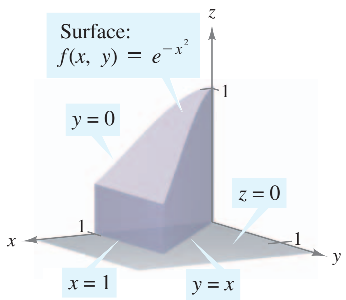
Figure 14.19

Try It

Exploration A Exploration B Exploration C Exploration D

In Examples 2 and 3, the problems could be solved with either order of integration because the regions were both vertically and horizontally simple. Moreover, had you used the order $dx\,dy$, you would have obtained integrals of comparable difficulty. There are, however, some occasions in which one order of integration is much more convenient than the other. Example 4 shows such a case.

EXAMPLE 4 Comparing Different Orders of Integration



Base is bounded by $y = 0$, $y = x$, and $x = 1$.
Figure 14.20

Rotatable Graph

Find the volume of the solid region R bounded by the surface

$$f(x, y) = e^{-x^2} \qquad \text{Surface}$$

and the planes $z = 0$, $y = 0$, $y = x$, and $x = 1$, as shown in Figure 14.20.

Solution The base of R in the xy-plane is bounded by the lines $y = 0$, $x = 1$, and $y = x$. The two possible orders of integration are shown in Figure 14.21.

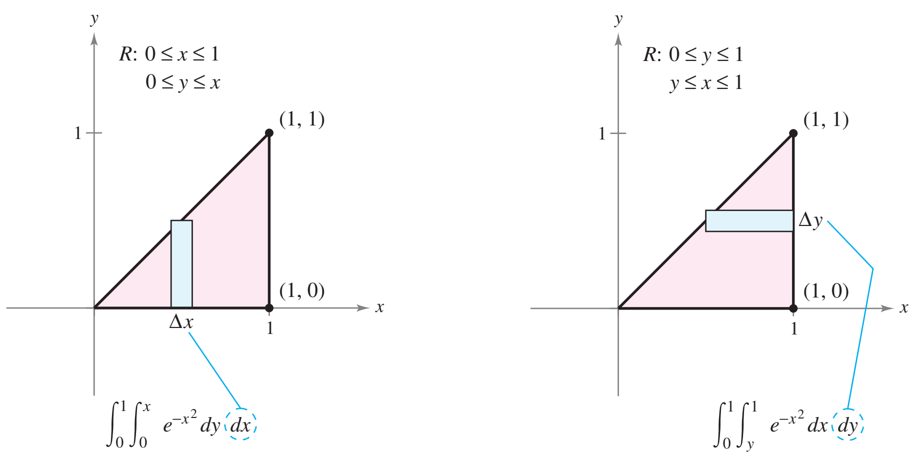


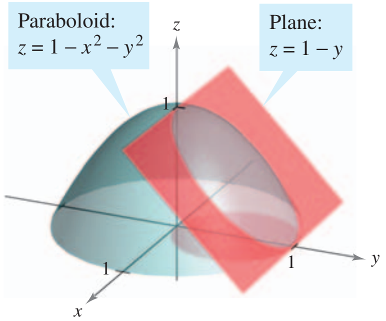
Figure 14.21

By setting up the corresponding iterated integrals, you can see that the order $dx\,dy$ requires the antiderivative $\int e^{-x^2}\,dx$, which is not an elementary function. On the other hand, the order $dy\,dx$ produces the integral

$$\begin{aligned}
\int_0^1 \int_0^x e^{-x^2}\,dy\,dx &= \int_0^1 e^{-x^2} y \Big]_0^x \,dx \\
&= \int_0^1 xe^{-x^2}\,dx \\
&= -\frac{1}{2}e^{-x^2}\Big]_0^1 \\
&= -\frac{1}{2}\left(\frac{1}{e} - 1\right) \\
&= \frac{e-1}{2e} \\
&\approx 0.316.
\end{aligned}$$

Try It | Exploration A | Open Exploration

NOTE Try using a symbolic integration utility to evaluate the integral in Example 4.



R: $0 \le y \le 1$
$-\sqrt{y-y^2} \le x \le \sqrt{y-y^2}$

Rotatable Graph

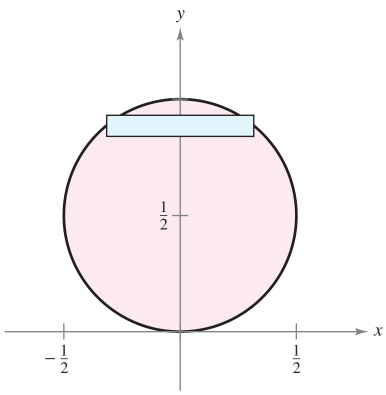


Figure 14.22

EXAMPLE 5 Volume of a Region Bounded by Two Surfaces

Find the volume of the solid region R bounded above by the paraboloid $z = 1 - x^2 - y^2$ and below by the plane $z = 1 - y$, as shown in Figure 14.22.

Solution Equating z-values, you can determine that the intersection of the two surfaces occurs on the right circular cylinder given by

$$1 - y = 1 - x^2 - y^2 \quad \Rightarrow \quad x^2 = y - y^2.$$

Because the volume of R is the difference between the volume under the paraboloid and the volume under the plane, you have

$$\begin{aligned}
\text{Volume} &= \int_0^1 \int_{-\sqrt{y-y^2}}^{\sqrt{y-y^2}} (1 - x^2 - y^2)\,dx\,dy - \int_0^1 \int_{-\sqrt{y-y^2}}^{\sqrt{y-y^2}} (1 - y)\,dx\,dy \\
&= \int_0^1 \int_{-\sqrt{y-y^2}}^{\sqrt{y-y^2}} (y - y^2 - x^2)\,dx\,dy \\
&= \int_0^1 \left[(y - y^2)x - \frac{x^3}{3} \right]_{-\sqrt{y-y^2}}^{\sqrt{y-y^2}} dy \\
&= \frac{4}{3} \int_0^1 (y - y^2)^{3/2}\,dy \\
&= \left(\frac{4}{3}\right)\left(\frac{1}{8}\right) \int_0^1 [1 - (2y - 1)^2]^{3/2}\,dy \\
&= \frac{1}{6} \int_{-\pi/2}^{\pi/2} \frac{\cos^4 \theta}{2}\,d\theta && \color{magenta}{2y - 1 = \sin \theta} \\
&= \frac{1}{6} \int_0^{\pi/2} \cos^4 \theta\,d\theta \\
&= \left(\frac{1}{6}\right)\left(\frac{3\pi}{16}\right) = \frac{\pi}{32}. && \color{magenta}{\text{Wallis's Formula}}
\end{aligned}$$

Try It **Exploration A**

Section 14.3 Change of Variables: Polar Coordinates

- **Write and evaluate double integrals in polar coordinates.**

Double Integrals in Polar Coordinates

Some double integrals are *much* easier to evaluate in polar form than in rectangular form. This is especially true for regions such as circles, cardioids, and rose curves, and for integrands that involve $x^2 + y^2$.

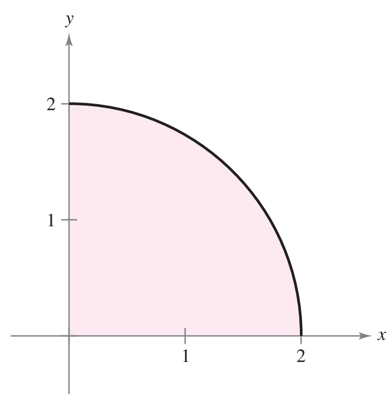
In Section 10.4, you learned that the polar coordinates (r, θ) of a point are related to the rectangular coordinates (x, y) of the point as follows.

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

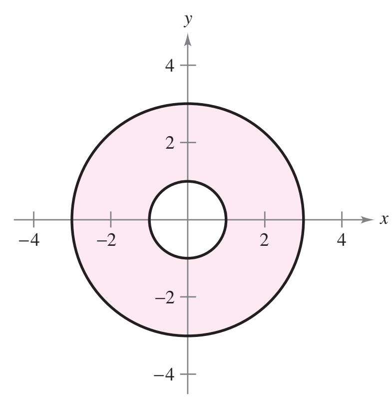
$$r^2 = x^2 + y^2 \quad \text{and} \quad \tan\theta = \frac{y}{x}$$

EXAMPLE 1 Using Polar Coordinates to Describe a Region

Use polar coordinates to describe each region shown in Figure 14.23.



(a)



(b)

Figure 14.23

Solution

a. The region R is a quarter circle of radius 2. It can be described in polar coordinates as

$$R = \{(r, \theta): 0 \le r \le 2, \quad 0 \le \theta \le \pi/2\}.$$

b. The region R consists of all points between the concentric circles of radii 1 and 3. It can be described in polar coordinates as

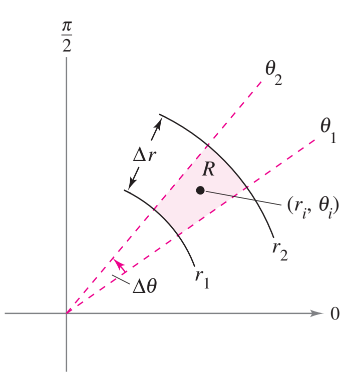
$$R = \{(r, \theta): 1 \le r \le 3, \quad 0 \le \theta \le 2\pi\}.$$

Try It **Exploration A**

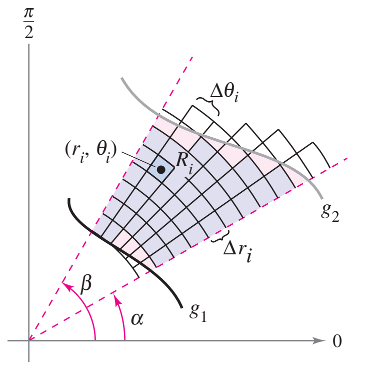
The regions in Example 1 are special cases of **polar sectors**

$$R = \{(r, \theta): r_1 \le r \le r_2, \quad \theta_1 \le \theta \le \theta_2\} \qquad \text{Polar sector}$$

as shown in Figure 14.24.



Polar sector
Figure 14.24



Polar grid is superimposed over region R.
Figure 14.25

To define a double integral of a continuous function $z = f(x, y)$ in polar coordinates, consider a region R bounded by the graphs of $r = g_1(\theta)$ and $r = g_2(\theta)$ and the lines $\theta = \alpha$ and $\theta = \beta$. Instead of partitioning R into small rectangles, use a partition of small polar sectors. On R, superimpose a polar grid made of rays and circular arcs, as shown in Figure 14.25. The polar sectors R_i lying entirely within R form an **inner polar partition** Δ, whose **norm** $\|\Delta\|$ is the length of the longest diagonal of the n polar sectors.

Consider a specific polar sector R_i, as shown in Figure 14.26. It can be shown (see Exercise 61) that the area of R_i is

$$\Delta A_i = r_i \Delta r_i \Delta\theta_i \qquad \text{Area of } R_i$$

where $\Delta r_i = r_2 - r_1$ and $\Delta\theta_i = \theta_2 - \theta_1$. This implies that the volume of the solid of height $f(r_i \cos\theta_i, r_i \sin\theta_i)$ above R_i is approximately

$$f(r_i \cos\theta_i, r_i \sin\theta_i) r_i \Delta r_i \Delta\theta_i$$

and you have

$$\int_R\!\!\int f(x, y)\, dA \approx \sum_{i=1}^{n} f(r_i \cos\theta_i, r_i \sin\theta_i) r_i \,\Delta\, r_i \,\Delta\theta_i.$$

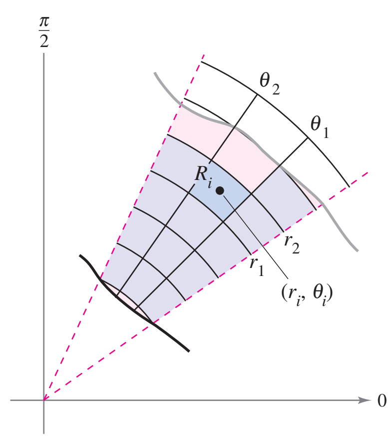
The sum on the right can be interpreted as a Riemann sum for $f(r\cos\theta, r\sin\theta)r$. The region R corresponds to a *horizontally simple* region S in the $r\theta$-plane, as shown in Figure 14.27. The polar sectors R_i correspond to rectangles S_i, and the area ΔA_i of S_i is $\Delta r_i \Delta\theta_i$. So, the right-hand side of the equation corresponds to the double integral

$$\int_S\!\!\int f(r\cos\theta, r\sin\theta) r\, dA.$$

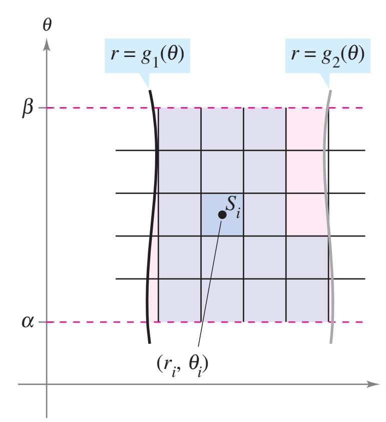
From this, you can apply Theorem 14.2 to write

$$\begin{aligned}\int_R\!\!\int f(x, y)\, dA &= \int_S\!\!\int f(r\cos\theta, r\sin\theta) r\, dA \\ &= \int_\alpha^\beta \int_{g_1(\theta)}^{g_2(\theta)} f(r\cos\theta, r\sin\theta) r\, dr\, d\theta.\end{aligned}$$

This suggests the following theorem, the proof of which is discussed in Section 14.8.



The polar sector R_i is the set of all points (r, θ) such that $r_1 \le r \le r_2$ and $\theta_1 \le \theta \le \theta_2$.
Figure 14.26



Horizontally simple region S
Figure 14.27

THEOREM 14.3 Change of Variables to Polar Form

Let R be a plane region consisting of all points $(x, y) = (r\cos\theta, r\sin\theta)$ satisfying the conditions $0 \le g_1(\theta) \le r \le g_2(\theta)$, $\alpha \le \theta \le \beta$, where $0 \le (\beta - \alpha) \le 2\pi$. If g_1 and g_2 are continuous on $[\alpha, \beta]$ and f is continuous on R, then

$$\int_R\!\!\int f(x, y)\, dA = \int_\alpha^\beta \int_{g_1(\theta)}^{g_2(\theta)} f(r\cos\theta, r\sin\theta) r\, dr\, d\theta.$$

EXPLORATION

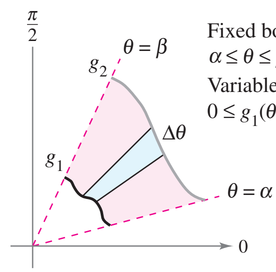
Volume of a Paraboloid Sector
In the Exploration feature on page 995, you were asked to summarize the different ways you know for finding the volume of the solid bounded by the paraboloid

$$z = a^2 - x^2 - y^2, \quad a > 0$$

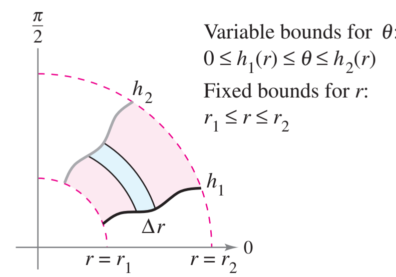
and the xy-plane. You now know another way. Use it to find the volume of the solid.

NOTE If $z = f(x, y)$ is nonnegative on R, then the integral in Theorem 14.3 can be interpreted as the volume of the solid region between the graph of f and the region R.

The region R is restricted to two basic types, ***r*-simple** regions and ***θ*-simple** regions, as shown in Figure 14.28.



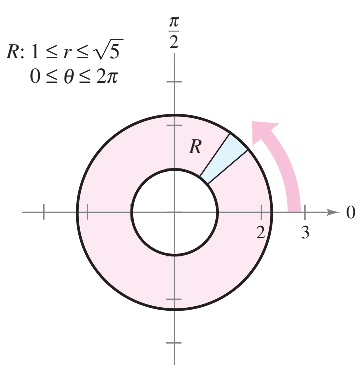
r-Simple region
Figure 14.28



θ-Simple region

EXAMPLE 2 Evaluating a Double Polar Integral

Let R be the annular region lying between the two circles $x^2 + y^2 = 1$ and $x^2 + y^2 = 5$. Evaluate the integral $\int_R\!\int (x^2 + y)\, dA$.



r-Simple region
Figure 14.29

Solution The polar boundaries are $1 \le r \le \sqrt{5}$ and $0 \le \theta \le 2\pi$, as shown in Figure 14.29. Furthermore, $x^2 = (r\cos\theta)^2$ and $y = r\sin\theta$. So, you have

$$\begin{aligned}
\int_R\!\!\int (x^2 + y)\, dA &= \int_0^{2\pi}\int_1^{\sqrt{5}} (r^2\cos^2\theta + r\sin\theta) r\, dr\, d\theta \\
&= \int_0^{2\pi}\int_1^{\sqrt{5}} (r^3\cos^2\theta + r^2\sin\theta)\, dr\, d\theta \\
&= \int_0^{2\pi} \left(\frac{r^4}{4}\cos^2\theta + \frac{r^3}{3}\sin\theta\right)\Big]_1^{\sqrt{5}}\, d\theta \\
&= \int_0^{2\pi} \left(6\cos^2\theta + \frac{5\sqrt{5} - 1}{3}\sin\theta\right) d\theta \\
&= \int_0^{2\pi} \left(3 + 3\cos 2\theta + \frac{5\sqrt{5} - 1}{3}\sin\theta\right) d\theta \\
&= \left(3\theta + \frac{3\sin 2\theta}{2} - \frac{5\sqrt{5} - 1}{3}\cos\theta\right)\Big]_0^{2\pi} \\
&= 6\pi.
\end{aligned}$$

Try It Exploration A Exploration B

In Example 2, be sure to notice the extra factor of r in the integrand. This comes from the formula for the area of a polar sector. In differential notation, you can write

$$dA = r\,dr\,d\theta$$

which indicates that the area of a polar sector increases as you move away from the origin.

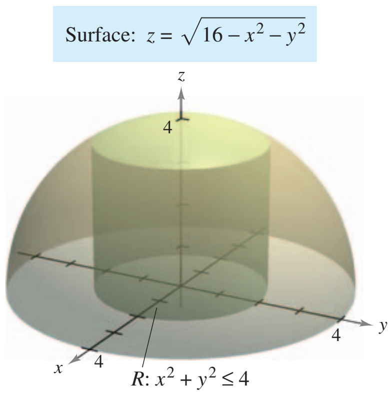


Figure 14.30

Rotatable Graph

EXAMPLE 3 Change of Variables to Polar Coordinates

Use polar coordinates to find the volume of the solid region bounded above by the hemisphere

$$z = \sqrt{16 - x^2 - y^2} \qquad \text{Hemisphere forms upper surface.}$$

and below by the circular region R given by

$$x^2 + y^2 \le 4 \qquad \text{Circular region forms lower surface.}$$

as shown in Figure 14.30.

Solution In Figure 14.30, you can see that R has the bounds

$$-\sqrt{4 - y^2} \le x \le \sqrt{4 - y^2}, \quad -2 \le y \le 2$$

and that $0 \le z \le \sqrt{16 - x^2 - y^2}$. In polar coordinates, the bounds are

$$0 \le r \le 2 \quad \text{and} \quad 0 \le \theta \le 2\pi$$

with height $z = \sqrt{16 - x^2 - y^2} = \sqrt{16 - r^2}$. Consequently, the volume V is given by

$$\begin{aligned} V = \int\!\!\int_R f(x, y)\,dA &= \int_0^{2\pi}\!\int_0^2 \sqrt{16 - r^2}\,r\,dr\,d\theta \\ &= -\frac{1}{3}\int_0^{2\pi} (16 - r^2)^{3/2}\Big]_0^2\,d\theta \\ &= -\frac{1}{3}\int_0^{2\pi} \left(24\sqrt{3} - 64\right)d\theta \\ &= -\frac{8}{3}\left(3\sqrt{3} - 8\right)\theta\Big]_0^{2\pi} \\ &= \frac{16\pi}{3}\left(8 - 3\sqrt{3}\right) \approx 46.979. \end{aligned}$$

Try It | Exploration A | Exploration B

NOTE To see the benefit of polar coordinates in Example 3, you should try to evaluate the corresponding rectangular iterated integral

$$\int_{-2}^{2}\int_{-\sqrt{4-y^2}}^{\sqrt{4-y^2}} \sqrt{16 - x^2 - y^2}\,dx\,dy.$$

TECHNOLOGY Any computer algebra system that can handle double integrals in rectangular coordinates can also handle double integrals in polar coordinates. The reason this is true is that once you have formed the iterated integral, its value is not changed by using different variables. In other words, if you use a computer algebra system to evaluate

$$\int_0^{2\pi}\!\int_0^2 \sqrt{16 - x^2}\,x\,dx\,dy$$

you should obtain the same value as that obtained in Example 3.

Just as with rectangular coordinates, the double integral

$$\int_R\int dA$$

can be used to find the area of a region in the plane.

EXAMPLE 4 Finding Areas of Polar Regions

Use a double integral to find the area enclosed by the graph of $r = 3\cos 3\theta$.

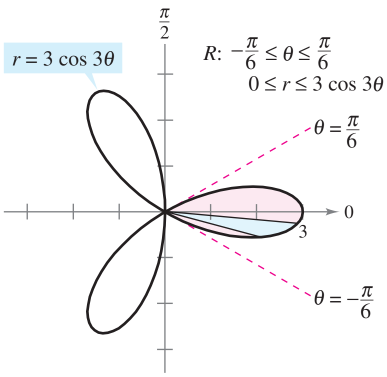


Figure 14.31

Editable Graph

Solution Let R be one petal of the curve shown in Figure 14.31. This region is r-simple, and the boundaries are as follows.

$$-\frac{\pi}{6} \le \theta \le \frac{\pi}{6} \qquad \text{Fixed bounds on } \theta$$

$$0 \le r \le 3\cos 3\theta \qquad \text{Variable bounds on } r$$

So, the area of one petal is

$$\begin{aligned}
\frac{1}{3}A = \int_R\int dA &= \int_{-\pi/6}^{\pi/6}\int_0^{3\cos 3\theta} r\,dr\,d\theta \\
&= \int_{-\pi/6}^{\pi/6} \left[\frac{r^2}{2}\right]_0^{3\cos 3\theta} d\theta \\
&= \frac{9}{2}\int_{-\pi/6}^{\pi/6} \cos^2 3\theta\,d\theta \\
&= \frac{9}{4}\int_{-\pi/6}^{\pi/6} (1 + \cos 6\theta)\,d\theta \\
&= \frac{9}{4}\left[\theta + \frac{1}{6}\sin 6\theta\right]_{-\pi/6}^{\pi/6} \\
&= \frac{3\pi}{4}.
\end{aligned}$$

So, the total area is $A = 9\pi/4$.

Try It | Exploration A | Exploration B | Open Exploration

As illustrated in Example 4, the area of a region in the plane can be represented by

$$A = \int_\alpha^\beta\int_{g_1(\theta)}^{g_2(\theta)} r\,dr\,d\theta.$$

If $g_1(\theta) = 0$, you obtain

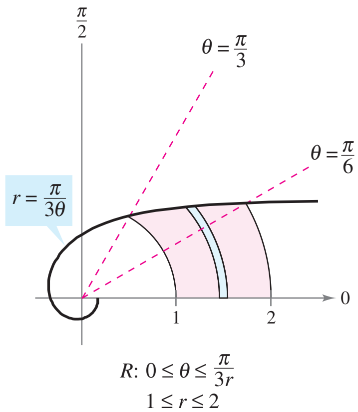
$$A = \int_\alpha^\beta\int_0^{g_2(\theta)} r\,dr\,d\theta = \int_\alpha^\beta \left[\frac{r^2}{2}\right]_0^{g_2(\theta)} d\theta = \int_\alpha^\beta \frac{1}{2}(g_2(\theta))^2\,d\theta$$

which agrees with Theorem 10.13.

So far in this section, all of the examples of iterated integrals in polar form have been of the form

$$\int_\alpha^\beta\int_{g_1(\theta)}^{g_2(\theta)} f(r\cos\theta, r\sin\theta)r\,dr\,d\theta$$

in which the order of integration is with respect to r first. Sometimes you can obtain a simpler integration problem by switching the order of integration, as illustrated in the next example.



θ-Simple region
Figure 14.32

Editable Graph

EXAMPLE 5 Changing the Order of Integration

Find the area of the region bounded above by the spiral

$$r = \frac{\pi}{3\theta}$$

and below by the polar axis, between $r = 1$ and $r = 2$.

Solution The region is shown in Figure 14.32. The polar boundaries for the region are

$$1 \le r \le 2 \quad \text{and} \quad 0 \le \theta \le \frac{\pi}{3r}.$$

So, the area of the region can be evaluated as follows.

$$A = \int_1^2 \int_0^{\pi/(3r)} r \, d\theta \, dr = \int_1^2 r\theta \Big]_0^{\pi/(3r)} dr = \int_1^2 \frac{\pi}{3} \, dr = \frac{\pi r}{3}\Big]_1^2 = \frac{\pi}{3}$$

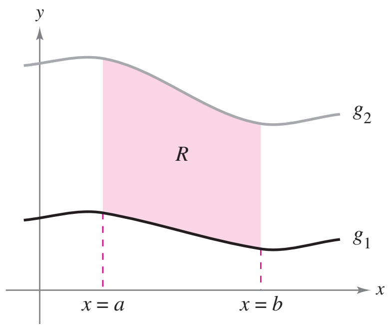
Try It

Exploration A

Section 14.4 Center of Mass and Moments of Inertia

- **Find the mass of a planar lamina using a double integral.**
- **Find the center of mass of a planar lamina using double integrals.**
- **Find moments of inertia using double integrals.**

Mass



Lamina of constant density ρ
Figure 14.33

Section 7.6 discussed several applications of integration involving a lamina of *constant* density ρ. For example, if the lamina corresponding to the region R, as shown in Figure 14.33, has a constant density ρ, then the mass of the lamina is given by

$$\text{Mass} = \rho A = \rho \int_R\int dA = \int_R\int \rho\, dA. \qquad \text{Constant density}$$

If not otherwise stated, a lamina is assumed to have a constant density. In this section, however, you will extend the definition of the term *lamina* to include thin plates of *variable* density. Double integrals can be used to find the mass of a lamina of *variable* density, where the density at (x, y) is given by the **density function ρ.**

Definition of Mass of a Planar Lamina of Variable Density

If ρ is a continuous density function on the lamina corresponding to a plane region R, then the mass m of the lamina is given by

$$m = \int_R\int \rho(x, y)\, dA. \qquad \text{Variable density}$$

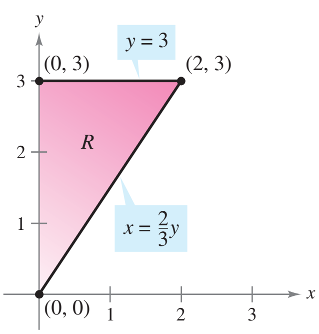
NOTE Density is normally expressed as mass per unit volume. For a planar lamina, however, density is mass per unit surface area.

EXAMPLE 1 Finding the Mass of a Planar Lamina

Find the mass of the triangular lamina with vertices $(0, 0)$, $(0, 3)$, and $(2, 3)$, given that the density at (x, y) is $\rho(x, y) = 2x + y$.

Solution As shown in Figure 14.34, region R has the boundaries $x = 0$, $y = 3$, and $y = 3x/2$ (or $x = 2y/3$). Therefore, the mass of the lamina is

$$\begin{aligned} m = \int_R\int (2x + y)\, dA &= \int_0^3 \int_0^{2y/3} (2x + y)\, dx\, dy \\ &= \int_0^3 \Big[x^2 + xy\Big]_0^{2y/3} dy \\ &= \frac{10}{9}\int_0^3 y^2\, dy \\ &= \frac{10}{9}\left[\frac{y^3}{3}\right]_0^3 \\ &= 10. \end{aligned}$$



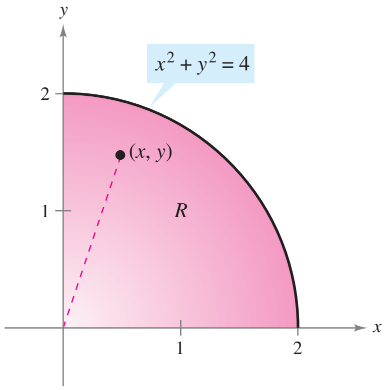
Lamina of variable density
$\rho(x, y) = 2x + y$
Figure 14.34

Editable Graph

Try It

Exploration A

NOTE In Figure 14.34, note that the planar lamina is shaded so that the darkest shading corresponds to the densest part.



Density at (x, y): $\rho(x, y) = k\sqrt{x^2 + y^2}$
Figure 14.35

Editable Graph

EXAMPLE 2 Finding Mass by Polar Coordinates

Find the mass of the lamina corresponding to the first-quadrant portion of the circle

$$x^2 + y^2 = 4$$

where the density at the point (x, y) is proportional to the distance between the point and the origin, as shown in Figure 14.35.

Solution At any point (x, y), the density of the lamina is

$$\begin{aligned}\rho(x, y) &= k\sqrt{(x-0)^2 + (y-0)^2}\\ &= k\sqrt{x^2 + y^2}.\end{aligned}$$

Because $0 \le x \le 2$ and $0 \le y \le \sqrt{4 - x^2}$, the mass is given by

$$\begin{aligned}m &= \int_R\int k\sqrt{x^2 + y^2}\, dA\\ &= \int_0^2 \int_0^{\sqrt{4-x^2}} k\sqrt{x^2 + y^2}\, dy\, dx.\end{aligned}$$

To simplify the integration, you can convert to polar coordinates, using the bounds $0 \le \theta \le \pi/2$ and $0 \le r \le 2$. So, the mass is

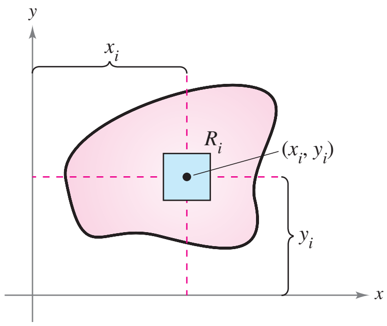
$$\begin{aligned}m = \int_R\int k\sqrt{x^2 + y^2}\, dA &= \int_0^{\pi/2}\int_0^2 k\sqrt{r^2}\, r\, dr\, d\theta\\ &= \int_0^{\pi/2}\int_0^2 kr^2\, dr\, d\theta\\ &= \int_0^{\pi/2} \frac{kr^3}{3}\Big]_0^2 d\theta\\ &= \frac{8k}{3}\int_0^{\pi/2} d\theta\\ &= \frac{8k}{3}\Big[\theta\Big]_0^{\pi/2}\\ &= \frac{4\pi k}{3}.\end{aligned}$$

Try It Exploration A

TECHNOLOGY On many occasions, this text has mentioned the benefits of computer programs that perform symbolic integration. Even if you use such a program regularly, you should remember that its greatest benefit comes only in the hands of a knowledgeable user. For instance, notice how much simpler the integral in Example 2 becomes when it is converted to polar form.

Rectangular Form	*Polar Form*
$\int_0^2\int_0^{\sqrt{4-x^2}} k\sqrt{x^2 + y^2}\, dy\, dx$	$\int_0^{\pi/2}\int_0^2 kr^2\, dr\, d\theta$

If you have access to software that performs symbolic integration, use it to evaluate both integrals. Some software programs cannot handle the first integral, but any program that can handle double integrals can evaluate the second integral.



$M_x = (\text{mass})(y_i)$
$M_y = (\text{mass})(x_i)$
Figure 14.36

Moments and Center of Mass

For a lamina of variable density, moments of mass are defined in a manner similar to that used for the uniform density case. For a partition Δ of a lamina corresponding to a plane region R, consider the ith rectangle R_i of one area ΔA_i, as shown in Figure 14.36. Assume that the mass of R_i is concentrated at one of its interior points (x_i, y_i). The moment of mass of R_i with respect to the x-axis can be approximated by

$$(\text{Mass})(y_i) \approx [\rho(x_i, y_i)\,\Delta A_i](y_i).$$

Similarly, the moment of mass with respect to the y-axis can be approximated by

$$(\text{Mass})(x_i) \approx [\rho(x_i, y_i)\,\Delta A_i](x_i).$$

By forming the Riemann sum of all such products and taking the limits as the norm of Δ approaches 0, you obtain the following definitions of moments of mass with respect to the x- and y-axes.

Moments and Center of Mass of a Variable Density Planar Lamina

Let ρ be a continuous density function on the planar lamina R. The **moments of mass** with respect to the x- and y-axes are

$$M_x = \int_R\!\!\int y\rho(x, y)\,dA \quad \text{and} \quad M_y = \int_R\!\!\int x\rho(x, y)\,dA.$$

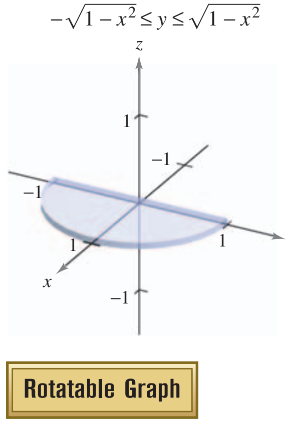
If m is the mass of the lamina, then the **center of mass** is

$$(\bar{x}, \bar{y}) = \left(\frac{M_y}{m}, \frac{M_x}{m}\right).$$

If R represents a simple plane region rather than a lamina, the point $(\bar{x}, \bar{y})$ is called the **centroid** of the region.

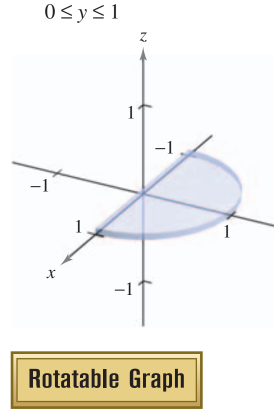
For some planar laminas with a constant density ρ, you can determine the center of mass (or one of its coordinates) using symmetry rather than using integration. For instance, consider the laminas of constant density shown in Figure 14.37. Using symmetry, you can see that $\bar{y} = 0$ for the first lamina and $\bar{x} = 0$ for the second lamina.

R: $0 \le x \le 1$
$-\sqrt{1-x^2} \le y \le \sqrt{1-x^2}$

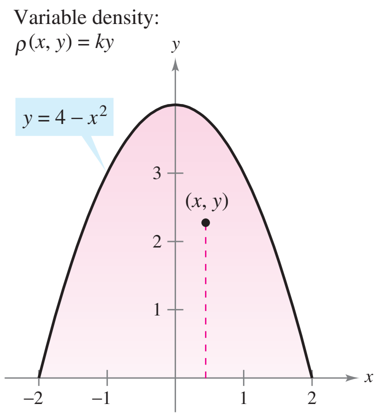


Lamina of constant density and symmetric with respect to the x-axis
Figure 14.37

R: $-\sqrt{1-y^2} \le x \le \sqrt{1-y^2}$
$0 \le y \le 1$



Lamina of constant density and symmetric with respect to the y-axis



The parabolic region of variable density
Figure 14.38

Editable Graph

EXAMPLE 3 Finding the Center of Mass

Find the center of mass of the lamina corresponding to the parabolic region

$$0 \le y \le 4 - x^2 \qquad \text{Parabolic region}$$

where the density at the point (x, y) is proportional to the distance between (x, y) and the x-axis, as shown in Figure 14.38.

Solution Because the lamina is symmetric with respect to the y-axis and

$$\rho(x, y) = ky$$

the center of mass lies on the y-axis. So, $\bar{x} = 0$. To find $\bar{y}$, first find the mass of the lamina.

$$\begin{aligned}
\text{Mass} = \int_{-2}^{2}\int_{0}^{4-x^2} ky\,dy\,dx &= \frac{k}{2}\int_{-2}^{2} y^2\Big]_{0}^{4-x^2} dx \\
&= \frac{k}{2}\int_{-2}^{2} (16 - 8x^2 + x^4)\,dx \\
&= \frac{k}{2}\left[16x - \frac{8x^3}{3} + \frac{x^5}{5}\right]_{-2}^{2} \\
&= k\left(32 - \frac{64}{3} + \frac{32}{5}\right) \\
&= \frac{256k}{15}
\end{aligned}$$

Next, find the moment about the x-axis.

$$\begin{aligned}
M_x = \int_{-2}^{2}\int_{0}^{4-x^2} (y)(ky)\,dy\,dx &= \frac{k}{3}\int_{-2}^{2} y^3\Big]_{0}^{4-x^2} dx \\
&= \frac{k}{3}\int_{-2}^{2} (64 - 48x^2 + 12x^4 - x^6)\,dx \\
&= \frac{k}{3}\left[64x - 16x^3 + \frac{12x^5}{5} - \frac{x^7}{7}\right]_{-2}^{2} \\
&= \frac{4096k}{105}
\end{aligned}$$

So,

$$\bar{y} = \frac{M_x}{m} = \frac{4096k/105}{256k/15} = \frac{16}{7}$$

and the center of mass is $\left(0, \frac{16}{7}\right)$.

Try It | Exploration A | Open Exploration

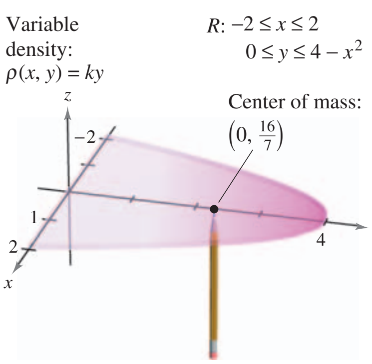


Figure 14.39

Rotatable Graph

Although you can think of the moments M_x and M_y as measuring the tendency to rotate about the x- or y-axis, the calculation of moments is usually an intermediate step toward a more tangible goal. The use of the moments M_x and M_y is typical—to find the center of mass. Determination of the center of mass is useful in a variety of applications that allow you to treat a lamina as if its mass were concentrated at just one point. Intuitively, you can think of the center of mass as the balancing point of the lamina. For instance, the lamina in Example 3 should balance on the point of a pencil placed at $\left(0, \frac{16}{7}\right)$, as shown in Figure 14.39.

Moments of Inertia

The moments of M_x and M_y used in determining the center of mass of a lamina are sometimes called the **first moments** about the x- and y-axes. In each case, the moment is the product of a mass times a distance.

$$M_x = \int_R\int \underbrace{(y)}_{\text{Distance to } x\text{-axis}}\underbrace{\rho(x, y)\, dA}_{\text{Mass}} \qquad M_y = \int_R\int \underbrace{(x)}_{\text{Distance to } y\text{-axis}}\underbrace{\rho(x, y)\, dA}_{\text{Mass}}$$

You will now look at another type of moment—the **second moment,** or the **moment of inertia** of a lamina about a line. In the same way that mass is a measure of the tendency of matter to resist a change in straight-line motion, the moment of inertia about a line is a *measure of the tendency of matter to resist a change in rotational motion*. For example, if a particle of mass m is a distance d from a fixed line, its moment of inertia about the line is defined as

$$I = md^2 = (\text{mass})(\text{distance})^2.$$

As with moments of mass, you can generalize this concept to obtain the moments of inertia about the x- and y-axes of a lamina of variable density. These second moments are denoted by I_x and I_y, and in each case the moment is the product of a mass times the square of a distance.

$$I_x = \int_R\int \underbrace{(y^2)}_{\text{Square of distance to } x\text{-axis}}\underbrace{\rho(x, y)\, dA}_{\text{Mass}} \qquad I_y = \int_R\int \underbrace{(x^2)}_{\text{Square of distance to } y\text{-axis}}\underbrace{\rho(x, y)\, dA}_{\text{Mass}}$$

The sum of the moments I_x and I_y is called the **polar moment of inertia** and is denoted by I_0.

NOTE For a lamina in the xy-plane, I_0 represents the moment of inertia of the lamina about the z-axis. The term "polar moment of inertia" stems from the fact that the square of the polar distance r is used in the calculation.

$$\begin{aligned} I_0 &= \int_R\int (x^2 + y^2)\rho(x, y)\, dA \\ &= \int_R\int r^2\rho(x, y)\, dA \end{aligned}$$

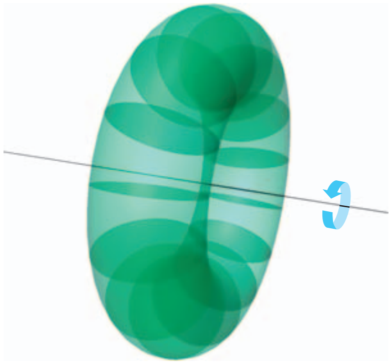
EXAMPLE 4 Finding the Moment of Inertia

Find the moment of inertia about the x-axis of the lamina in Example 3.

Solution From the definition of moment of inertia, you have

$$\begin{aligned} I_x &= \int_{-2}^{2}\int_{0}^{4-x^2} y^2(ky)\, dy\, dx \\ &= \frac{k}{4}\int_{-2}^{2} y^4\Big]_0^{4-x^2} dx \\ &= \frac{k}{4}\int_{-2}^{2} (256 - 256x^2 + 96x^4 - 16x^6 + x^8)\, dx \\ &= \frac{k}{4}\left[256x - \frac{256x^3}{3} + \frac{96x^5}{5} - \frac{16x^7}{7} + \frac{x^9}{9}\right]_{-2}^{2} \\ &= \frac{32{,}768k}{315}. \end{aligned}$$

Try It Exploration A



Planar lamina revolving at ω radians per second
Figure 14.40

Rotatable Graph

The moment of inertia I of a revolving lamina can be used to measure its kinetic energy. For example, suppose a planar lamina is revolving about a line with an **angular speed** of ω radians per second, as shown in Figure 14.40. The kinetic energy E of the revolving lamina is

$$E = \frac{1}{2} I\omega^2. \qquad \text{Kinetic energy for rotational motion}$$

On the other hand, the kinetic energy E of a mass m moving in a straight line at a velocity v is

$$E = \frac{1}{2} mv^2. \qquad \text{Kinetic energy for linear motion}$$

So, the kinetic energy of a mass moving in a straight line is proportional to its mass, but the kinetic energy of a mass revolving about an axis is proportional to its moment of inertia.

The **radius of gyration** $\bar{\bar{r}}$ of a revolving mass m with moment of inertia I is defined to be

$$\bar{\bar{r}} = \sqrt{\frac{I}{m}}. \qquad \text{Radius of gyration}$$

If the entire mass were located at a distance $\bar{\bar{r}}$ from its axis of revolution, it would have the same moment of inertia and, consequently, the same kinetic energy. For instance, the radius of gyration of the lamina in Example 4 about the x-axis is given by

$$\bar{\bar{y}} = \sqrt{\frac{I_x}{m}} = \sqrt{\frac{32{,}768k/315}{256k/15}} = \sqrt{\frac{128}{21}} \approx 2.469.$$

EXAMPLE 5 Finding the Radius of Gyration

Find the radius of gyration about the y-axis for the lamina corresponding to the region R: $0 \le y \le \sin x$, $0 \le x \le \pi$, where the density at (x, y) is given by $\rho(x, y) = x$.

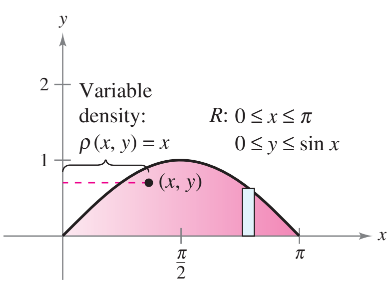


Figure 14.41

Editable Graph

Solution The region R is shown in Figure 14.41. By integrating $\rho(x, y) = x$ over the region R, you can determine that the mass of the region is π. The moment of inertia about the y-axis is

$$\begin{aligned}
I_y &= \int_0^{\pi}\int_0^{\sin x} x^3\,dy\,dx \\
&= \int_0^{\pi} x^3 y\Big]_0^{\sin x} dx \\
&= \int_0^{\pi} x^3 \sin x\,dx \\
&= \Big[(3x^2 - 6)(\sin x) - (x^3 - 6x)(\cos x)\Big]_0^{\pi} \\
&= \pi^3 - 6\pi.
\end{aligned}$$

So, the radius of gyration about the y-axis is

$$\begin{aligned}
\bar{\bar{x}} &= \sqrt{\frac{I_y}{m}} \\
&= \sqrt{\frac{\pi^3 - 6\pi}{\pi}} \\
&= \sqrt{\pi^2 - 6} \approx 1.967.
\end{aligned}$$

Try It Exploration A

Section 14.5 Surface Area

- **Use a double integral to find the area of a surface.**

Surface Area

At this point you know a great deal about the solid region lying between a surface and a closed and bounded region R in the xy-plane, as shown in Figure 14.42. For example, you know how to find the extrema of f on R (Section 13.8), the area of the base R of the solid (Section 14.1), the volume of the solid (Section 14.2), and the centroid of the base R (Section 14.4).

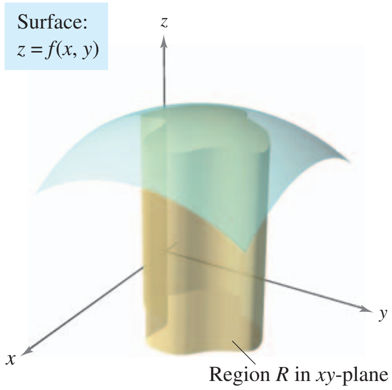


Figure 14.42

Rotatable Graph

In this section, you will learn how to find the upper **surface area** of the solid. Later, you will learn how to find the centroid of the solid (Section 14.6) and the lateral surface area (Section 15.2).

To begin, consider a surface S given by

$$z = f(x, y) \qquad \text{Surface defined over a region } R$$

defined over a region R. Assume that R is closed and bounded and that f has continuous first partial derivatives. To find the surface area, construct an inner partition of R consisting of n rectangles, where the area of the ith rectangle R_i is $\Delta A_i = \Delta x_i \, \Delta y_i$, as shown in Figure 14.43. In each R_i let (x_i, y_i) be the point that is closest to the origin. At the point $(x_i, y_i, z_i) = (x_i, y_i, f(x_i, y_i))$ on the surface S, construct a tangent plane T_i. The area of the portion of the tangent plane that lies directly above R_i is approximately equal to the area of the surface lying directly above R_i. That is, $\Delta T_i \approx \Delta S_i$. So, the surface area of S is given by

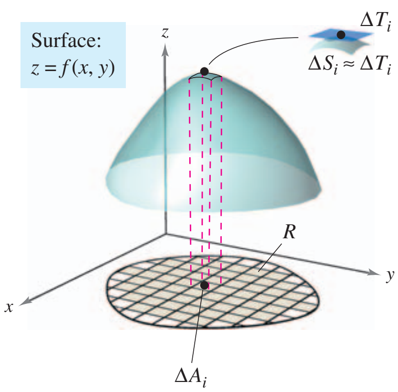


Figure 14.43

Rotatable Graph

$$\sum_{i=1}^{n} \Delta S_i \approx \sum_{i=1}^{n} \Delta T_i.$$

To find the area of the parallelogram ΔT_i, note that its sides are given by the vectors

$$\mathbf{u} = \Delta x_i \mathbf{i} + f_x(x_i, y_i)\,\Delta x_i \mathbf{k}$$

and

$$\mathbf{v} = \Delta y_i \mathbf{j} + f_y(x_i, y_i)\,\Delta y_i \mathbf{k}.$$

From Theorem 11.8, the area of ΔT_i is given by $\|\mathbf{u} \times \mathbf{v}\|$, where

$$\begin{aligned}
\mathbf{u} \times \mathbf{v} &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \Delta x_i & 0 & f_x(x_i, y_i)\,\Delta x_i \\ 0 & \Delta y_i & f_y(x_i, y_i)\,\Delta y_i \end{vmatrix} \\
&= -f_x(x_i, y_i)\,\Delta x_i \Delta y_i \mathbf{i} - f_y(x_i, y_i)\,\Delta x_i \Delta y_i \mathbf{j} + \Delta x_i \Delta y_i \mathbf{k} \\
&= (-f_x(x_i, y_i)\mathbf{i} - f_y(x_i, y_i)\mathbf{j} + \mathbf{k})\,\Delta A_i.
\end{aligned}$$

So, the area of ΔT_i is $\|\mathbf{u} \times \mathbf{v}\| = \sqrt{[f_x(x_i, y_i)]^2 + [f_y(x_i, y_i)]^2 + 1}\,\Delta A_i$, and

$$\begin{aligned}
\text{Surface area of } S &\approx \sum_{i=1}^{n} \Delta S_i \\
&\approx \sum_{i=1}^{n} \sqrt{1 + [f_x(x_i, y_i)]^2 + [f_y(x_i, y_i)]^2}\,\Delta A_i.
\end{aligned}$$

This suggests the following definition of surface area.

Definition of Surface Area

If f and its first partial derivatives are continuous on the closed region R in the xy-plane, then the **area of the surface S** given by $z = f(x, y)$ over R is given by

$$\text{Surface area} = \iint_R dS$$
$$= \iint_R \sqrt{1 + [f_x(x, y)]^2 + [f_y(x, y)]^2}\, dA.$$

As an aid to remembering the double integral for surface area, it is helpful to note its similarity to the integral for arc length.

Length on x-axis: $\displaystyle\int_a^b dx$

Arc length in xy-plane: $\displaystyle\int_a^b ds = \int_a^b \sqrt{1 + [f'(x)]^2}\, dx$

Area in xy-plane: $\displaystyle\iint_R dA$

Surface area in space: $\displaystyle\iint_R dS = \iint_R \sqrt{1 + [f_x(x, y)]^2 + [f_y(x, y)]^2}\, dA$

Like integrals for arc length, integrals for surface area are often very difficult to evaluate. However, one type that is easily evaluated is demonstrated in the next example.

EXAMPLE 1 The Surface Area of a Plane Region

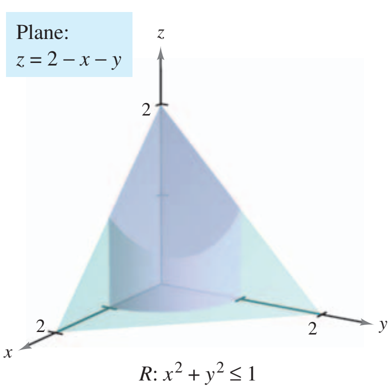


Figure 14.44

Rotatable Graph

Find the surface area of the portion of the plane

$$z = 2 - x - y$$

that lies above the circle $x^2 + y^2 \le 1$ in the first quadrant, as shown in Figure 14.44.

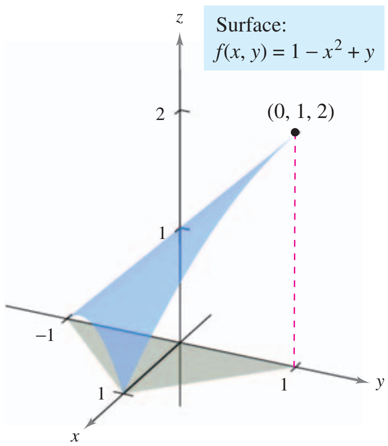
Solution Because $f_x(x, y) = -1$ and $f_y(x, y) = -1$, the surface area is given by

$$S = \iint_R \sqrt{1 + [f_x(x, y)]^2 + [f_y(x, y)]^2}\, dA \qquad \text{Formula for surface area}$$
$$= \iint_R \sqrt{1 + (-1)^2 + (-1)^2}\, dA \qquad \text{Substitute.}$$
$$= \iint_R \sqrt{3}\, dA$$
$$= \sqrt{3} \iint_R dA.$$

Note that the last integral is simply $\sqrt{3}$ times the area of the region R. R is a quarter circle of radius 1, with an area of $\frac{1}{4}\pi(1^2)$ or $\pi/4$. So, the area of S is

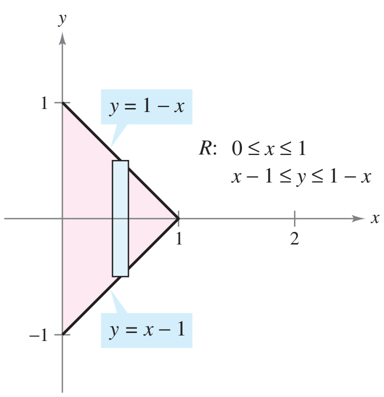
$$S = \sqrt{3}\,(\text{area of } R)$$
$$= \sqrt{3}\left(\frac{\pi}{4}\right)$$
$$= \frac{\sqrt{3}\,\pi}{4}.$$

Try It Exploration A



(a)

Rotatable Graph



(b)
Figure 14.45

EXAMPLE 2 Finding Surface Area

Find the area of the portion of the surface

$$f(x, y) = 1 - x^2 + y$$

that lies above the triangular region with vertices (1, 0, 0), (0, −1, 0), and (0, 1, 0), as shown in Figure 14.45(a).

Solution Because $f_x(x, y) = -2x$ and $f_y(x, y) = 1$, you have

$$S = \iint_R \sqrt{1 + [f_x(x, y)]^2 + [f_y(x, y)]^2}\, dA = \iint_R \sqrt{1 + 4x^2 + 1}\, dA.$$

In Figure 14.45(b), you can see that the bounds for R are $0 \le x \le 1$ and $x - 1 \le y \le 1 - x$. So, the integral becomes

$$\begin{aligned} S &= \int_0^1 \int_{x-1}^{1-x} \sqrt{2 + 4x^2}\, dy\, dx \\ &= \int_0^1 y\sqrt{2 + 4x^2}\Big]_{x-1}^{1-x} dx \\ &= \int_0^1 \Big[(1 - x)\sqrt{2 + 4x^2} - (x - 1)\sqrt{2 + 4x^2}\Big] dx \\ &= \int_0^1 \left(2\sqrt{2 + 4x^2} - 2x\sqrt{2 + 4x^2}\right) dx \qquad \text{Integration tables (Appendix B), Formula 26 and Power Rule} \\ &= \left[x\sqrt{2 + 4x^2} + \ln\left(2x + \sqrt{2 + 4x^2}\right) - \frac{(2 + 4x^2)^{3/2}}{6}\right]_0^1 \\ &= \sqrt{6} + \ln\left(2 + \sqrt{6}\right) - \sqrt{6} - \ln\sqrt{2} + \frac{1}{3}\sqrt{2} \approx 1.618. \end{aligned}$$

Try It   Exploration A   Open Exploration

EXAMPLE 3 Change of Variables to Polar Coordinates

Find the surface area of the paraboloid $z = 1 + x^2 + y^2$ that lies above the unit circle, as shown in Figure 14.46.

Solution Because $f_x(x, y) = 2x$ and $f_y(x, y) = 2y$, you have

$$S = \iint_R \sqrt{1 + [f_x(x, y)]^2 + [f_y(x, y)]^2}\, dA = \iint_R \sqrt{1 + 4x^2 + 4y^2}\, dA.$$

You can convert to polar coordinates by letting $x = r\cos\theta$ and $y = r\sin\theta$. Then, because the region R is bounded by $0 \le r \le 1$ and $0 \le \theta \le 2\pi$, you have

$$\begin{aligned} S &= \int_0^{2\pi} \int_0^1 \sqrt{1 + 4r^2}\, r\, dr\, d\theta \\ &= \int_0^{2\pi} \frac{1}{12}(1 + 4r^2)^{3/2}\Big]_0^1 d\theta \\ &= \int_0^{2\pi} \frac{5\sqrt{5} - 1}{12}\, d\theta \\ &= \frac{5\sqrt{5} - 1}{12}\,\theta\Big]_0^{2\pi} \\ &= \frac{\pi\left(5\sqrt{5} - 1\right)}{6} \\ &\approx 5.33. \end{aligned}$$

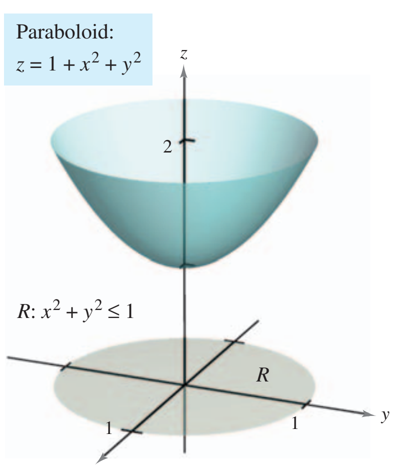


Figure 14.46

Rotatable Graph

Try It   Exploration A

EXAMPLE 4 Finding Surface Area

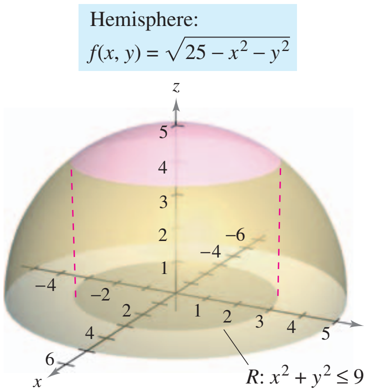


Figure 14.47

Rotatable Graph

Find the surface area S of the portion of the hemisphere

$$f(x, y) = \sqrt{25 - x^2 - y^2} \qquad \text{Hemisphere}$$

that lies above the region R bounded by the circle $x^2 + y^2 \leq 9$, as shown in Figure 14.47.

Solution The first partial derivatives of f are

$$f_x(x, y) = \frac{-x}{\sqrt{25 - x^2 - y^2}} \quad \text{and} \quad f_y(x, y) = \frac{-y}{\sqrt{25 - x^2 - y^2}}$$

and, from the formula for surface area, you have

$$\begin{aligned} dS &= \sqrt{1 + [f_x(x, y)]^2 + [f_y(x, y)]^2}\, dA \\ &= \sqrt{1 + \left(\frac{-x}{\sqrt{25 - x^2 - y^2}}\right)^2 + \left(\frac{-y}{\sqrt{25 - x^2 - y^2}}\right)^2}\, dA \\ &= \frac{5}{\sqrt{25 - x^2 - y^2}}\, dA. \end{aligned}$$

So, the surface area is

$$S = \iint_R \frac{5}{\sqrt{25 - x^2 - y^2}}\, dA.$$

You can convert to polar coordinates by letting $x = r \cos \theta$ and $y = r \sin \theta$. Then, because the region R is bounded by $0 \leq r \leq 3$ and $0 \leq \theta \leq 2\pi$, you obtain

$$\begin{aligned} S &= \int_0^{2\pi} \int_0^3 \frac{5}{\sqrt{25 - r^2}}\, r\, dr\, d\theta \\ &= 5 \int_0^{2\pi} -\sqrt{25 - r^2}\,\Big]_0^3\, d\theta \\ &= 5 \int_0^{2\pi} d\theta \\ &= 10\pi. \end{aligned}$$

Try It Exploration A

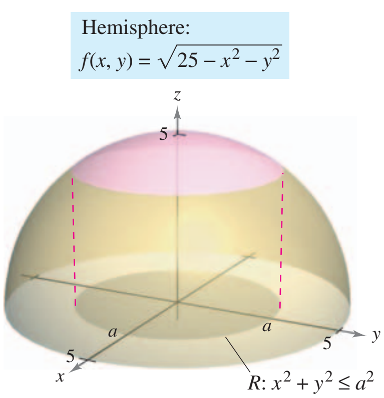


Figure 14.48

Rotatable Graph

The procedure used in Example 4 can be extended to find the surface area of a sphere by using the region R bounded by the circle $x^2 + y^2 \leq a^2$, where $0 < a < 5$, as shown in Figure 14.48. The surface area of the portion of the hemisphere

$$f(x, y) = \sqrt{25 - x^2 - y^2}$$

lying above the circular region can be shown to be

$$\begin{aligned} S &= \iint_R \frac{5}{\sqrt{25 - x^2 - y^2}}\, dA \\ &= \int_0^{2\pi} \int_0^a \frac{5}{\sqrt{25 - r^2}}\, r\, dr\, d\theta \\ &= 10\pi\left(5 - \sqrt{25 - a^2}\right). \end{aligned}$$

By taking the limit as a approaches 5 and doubling the result, you obtain a total area of 100π. (The surface area of a sphere of radius r is $S = 4\pi r^2$.)

You can use Simpson's Rule or the Trapezoidal Rule to approximate the value of a double integral, *provided* you can get through the first integration. This is demonstrated in the next example.

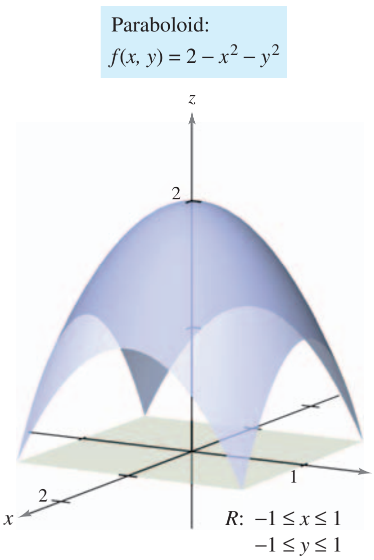


Figure 14.49

Rotatable Graph

EXAMPLE 5 Approximating Surface Area by Simpson's Rule

Find the area of the surface of the paraboloid

$$f(x, y) = 2 - x^2 - y^2 \qquad \text{Paraboloid}$$

that lies above the square region bounded by $-1 \le x \le 1$ and $-1 \le y \le 1$, as shown in Figure 14.49.

Solution Using the partial derivatives

$$f_x(x, y) = -2x \quad \text{and} \quad f_y(x, y) = -2y$$

you have a surface area of

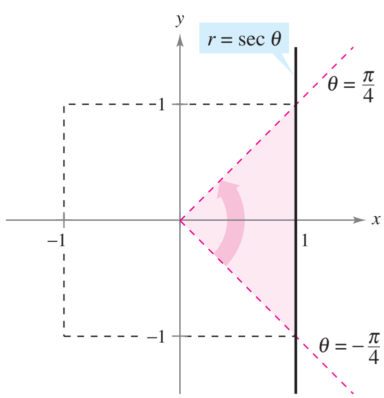
$$\begin{aligned} S &= \int\!\!\int_R \sqrt{1 + [f_x(x, y)]^2 + [f_y(x, y)]^2}\, dA \\ &= \int\!\!\int_R \sqrt{1 + (-2x)^2 + (-2y)^2}\, dA \\ &= \int\!\!\int_R \sqrt{1 + 4x^2 + 4y^2}\, dA. \end{aligned}$$

In polar coordinates, the line $x = 1$ is given by $r \cos \theta = 1$ or $r = \sec \theta$, and you can determine from Figure 14.50 that one-fourth of the region R is bounded by

$$0 \le r \le \sec \theta \quad \text{and} \quad -\frac{\pi}{4} \le \theta \le \frac{\pi}{4}.$$

Letting $x = r \cos \theta$ and $y = r \sin \theta$ produces

$$\begin{aligned} \frac{1}{4} S &= \frac{1}{4} \int\!\!\int_R \sqrt{1 + 4x^2 + 4y^2}\, dA \\ &= \int_{-\pi/4}^{\pi/4} \int_0^{\sec \theta} \sqrt{1 + 4r^2}\, r\, dr\, d\theta \\ &= \int_{-\pi/4}^{\pi/4} \frac{1}{12}(1 + 4r^2)^{3/2} \Big]_0^{\sec \theta} d\theta \\ &= \frac{1}{12} \int_{-\pi/4}^{\pi/4} \left[(1 + 4 \sec^2 \theta)^{3/2} - 1\right] d\theta. \end{aligned}$$



One-fourth of the region R is bounded by $0 \le r \le \sec \theta$ and $-\frac{\pi}{4} \le \theta \le \frac{\pi}{4}$.

Figure 14.50

Finally, using Simpson's Rule with $n = 10$, you can approximate this single integral to be

$$\begin{aligned} S &= \frac{1}{3} \int_{-\pi/4}^{\pi/4} \left[(1 + 4 \sec^2 \theta)^{3/2} - 1\right] d\theta \\ &\approx 7.450. \end{aligned}$$

Try It   Exploration A

TECHNOLOGY Most computer programs that are capable of performing symbolic integration for multiple integrals are also capable of performing numerical approximation techniques. If you have access to such software, use it to approximate the value of the integral in Example 5.

Section 14.6 Triple Integrals and Applications

- **Use a triple integral to find the volume of a solid region.**
- **Find the center of mass and moments of inertia of a solid region.**

Triple Integrals

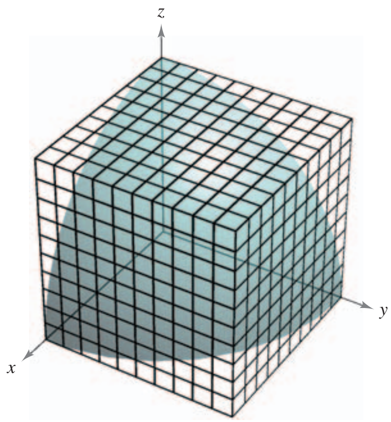
The procedure used to define a **triple integral** follows that used for double integrals. Consider a function f of three variables that is continuous over a bounded solid region Q. Then, encompass Q with a network of boxes and form the **inner partition** consisting of all boxes lying entirely within Q, as shown in Figure 14.51. The volume of the ith box is

$$\Delta V_i = \Delta x_i \Delta y_i \Delta z_i. \qquad \text{Volume of } i\text{th box}$$

The **norm** $\|\Delta\|$ of the partition is the length of the longest diagonal of the n boxes in the partition. Choose a point (x_i, y_i, z_i) in each box and form the Riemann sum

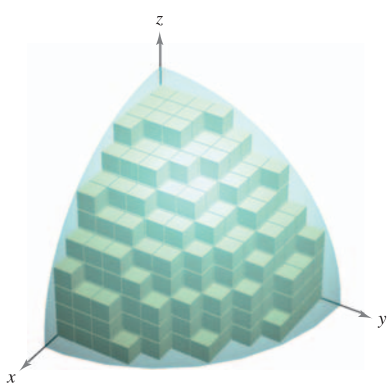
$$\sum_{i=1}^{n} f(x_i, y_i, z_i)\,\Delta V_i.$$

Taking the limit as $\|\Delta\| \to 0$ leads to the following definition.



Solid region Q

Rotatable Graph



Volume of $Q \approx \sum_{i=1}^{n} f(x_i, y_i, z_i)\,\Delta V_i$

Figure 14.51

Rotatable Graph

Definition of Triple Integral

If f is continuous over a bounded solid region Q, then the **triple integral of f over Q** is defined as

$$\iiint_Q f(x, y, z)\,dV = \lim_{\|\Delta\| \to 0} \sum_{i=1}^{n} f(x_i, y_i, z_i)\,\Delta V_i$$

provided the limit exists. The **volume** of the solid region Q is given by

$$\text{Volume of } Q = \iiint_Q dV.$$

Some of the properties of double integrals in Theorem 14.1 can be restated in terms of triple integrals.

1. $\displaystyle \iiint_Q cf(x, y, z)\,dV = c\iiint_Q f(x, y, z)\,dV$

2. $\displaystyle \iiint_Q [f(x, y, z) \pm g(x, y, z)]\,dV = \iiint_Q f(x, y, z)\,dV \pm \iiint_Q g(x, y, z)\,dV$

3. $\displaystyle \iiint_Q f(x, y, z)\,dV = \iiint_{Q_1} f(x, y, z)\,dV + \iiint_{Q_2} f(x, y, z)\,dV$

In the properties above, Q is the union of two nonoverlapping solid subregions Q_1 and Q_2. If the solid region Q is simple, the triple integral $\iiint f(x, y, z)\,dV$ can be evaluated with an iterated integral using one of the six possible orders of integration:

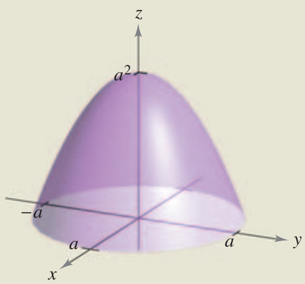
$$dx\,dy\,dz \quad dy\,dx\,dz \quad dz\,dx\,dy \quad dx\,dz\,dy \quad dy\,dz\,dx \quad dz\,dy\,dx.$$

EXPLORATION

Volume of a Paraboloid Sector On pages 995 and 1003, you were asked to summarize the different ways you know for finding the volume of the solid bounded by the paraboloid

$$z = a^2 - x^2 - y^2, \quad a > 0$$

and the xy-plane. You now know one more way. Use it to find the volume of the solid.



Rotatable Graph

The following version of Fubini's Theorem describes a region that is considered simple with respect to the order $dz\,dy\,dx$. Similar descriptions can be given for the other five orders.

THEOREM 14.4 Evaluation by Iterated Integrals

Let f be continuous on a solid region Q defined by

$$a \le x \le b, \quad h_1(x) \le y \le h_2(x), \quad g_1(x, y) \le z \le g_2(x, y)$$

where h_1, h_2, g_1, and g_2 are continuous functions. Then,

$$\iiint_Q f(x, y, z)\, dV = \int_a^b \int_{h_1(x)}^{h_2(x)} \int_{g_1(x, y)}^{g_2(x, y)} f(x, y, z)\, dz\, dy\, dx.$$

To evaluate a triple iterated integral in the order $dz\,dy\,dx$, hold *both* x and y constant for the innermost integration. Then, hold x constant for the second integration.

EXAMPLE 1 Evaluating a Triple Iterated Integral

Evaluate the triple iterated integral

$$\int_0^2 \int_0^x \int_0^{x+y} e^x(y + 2z)\, dz\, dy\, dx.$$

Solution For the first integration, hold x and y constant and integrate with respect to z.

$$\begin{aligned}\int_0^2 \int_0^x \int_0^{x+y} e^x(y + 2z)\, dz\, dy\, dx &= \int_0^2 \int_0^x e^x(yz + z^2)\Big]_0^{x+y} dy\, dx \\ &= \int_0^2 \int_0^x e^x(x^2 + 3xy + 2y^2)\, dy\, dx\end{aligned}$$

For the second integration, hold x constant and integrate with respect to y.

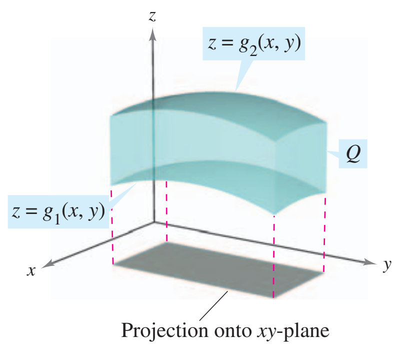
$$\begin{aligned}\int_0^2 \int_0^x e^x(x^2 + 3xy + 2y^2)\, dy\, dx &= \int_0^2 \left[e^x\left(x^2y + \frac{3xy^2}{2} + \frac{2y^3}{3}\right)\right]_0^x dx \\ &= \frac{19}{6}\int_0^2 x^3e^x\, dx\end{aligned}$$

Finally, integrate with respect to x.

$$\begin{aligned}\frac{19}{6}\int_0^2 x^3e^x\, dx &= \frac{19}{6}\Big[e^x(x^3 - 3x^2 + 6x - 6)\Big]_0^2 \\ &= 19\left(\frac{e^2}{3} + 1\right) \\ &\approx 65.797\end{aligned}$$

Try It Exploration A Exploration B

Example 1 demonstrates the integration order $dz\,dy\,dx$. For other orders, you can follow a similar procedure. For instance, to evaluate a triple iterated integral in the order $dx\,dy\,dz$, hold both y and z constant for the innermost integration and integrate with respect to x. Then, for the second integration, hold z constant and integrate with respect to y. Finally, for the third integration, integrate with respect to z.



Solid region Q lies between two surfaces.
Figure 14.52

Rotatable Graph

To find the limits for a particular order of integration, it is generally advisable first to determine the innermost limits, which may be functions of the outer two variables. Then, by projecting the solid Q onto the coordinate plane of the outer two variables, you can determine their limits of integration by the methods used for double integrals. For instance, to evaluate

$$\iiint_Q f(x, y, z)\, dz\, dy\, dx$$

first determine the limits for z, and then the integral has the form

$$\iint \left[\int_{g_1(x, y)}^{g_2(x, y)} f(x, y, z)\, dz \right] dy\, dx.$$

By projecting the solid Q onto the xy-plane, you can determine the limits for x and y as you did for double integrals, as shown in Figure 14.52.

EXAMPLE 2 Using a Triple Integral to Find Volume

Find the volume of the ellipsoid given by $4x^2 + 4y^2 + z^2 = 16$.

Solution Because x, y, and z play similar roles in the equation, the order of integration is probably immaterial, and you can arbitrarily choose $dz\, dy\, dx$. Moreover, you can simplify the calculation by considering only the portion of the ellipsoid lying in the first octant, as shown in Figure 14.53. From the order $dz\, dy\, dx$, you first determine the bounds for z.

$$0 \le z \le 2\sqrt{4 - x^2 - y^2}$$

In Figure 14.54, you can see that the boundaries for x and y are $0 \le x \le 2$ and $0 \le y \le \sqrt{4 - x^2}$, so the volume of the ellipsoid is

$$\begin{aligned}
V &= \iiint_Q dV \\
&= 8 \int_0^2 \int_0^{\sqrt{4-x^2}} \int_0^{2\sqrt{4-x^2-y^2}} dz\, dy\, dx \\
&= 8 \int_0^2 \int_0^{\sqrt{4-x^2}} z \Big]_0^{2\sqrt{4-x^2-y^2}} dy\, dx \\
&= 16 \int_0^2 \int_0^{\sqrt{4-x^2}} \sqrt{(4 - x^2) - y^2}\, dy\, dx \qquad \text{Integration tables (Appendix B), Formula 37} \\
&= 8 \int_0^2 \left[y\sqrt{4 - x^2 - y^2} + (4 - x^2) \arcsin\left(\frac{y}{\sqrt{4 - x^2}}\right) \right]_0^{\sqrt{4-x^2}} dx \\
&= 8 \int_0^2 [0 + (4 - x^2)\arcsin(1) - 0 - 0]\, dx \\
&= 8 \int_0^2 (4 - x^2)\left(\frac{\pi}{2}\right) dx \\
&= 4\pi \left[4x - \frac{x^3}{3} \right]_0^2 \\
&= \frac{64\pi}{3}.
\end{aligned}$$

Try It Exploration A Exploration B

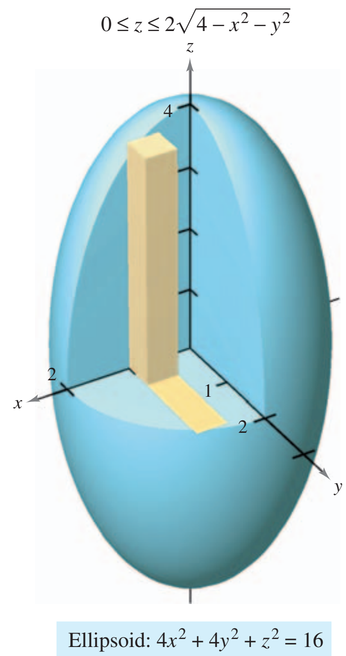


Figure 14.53

Rotatable Graph

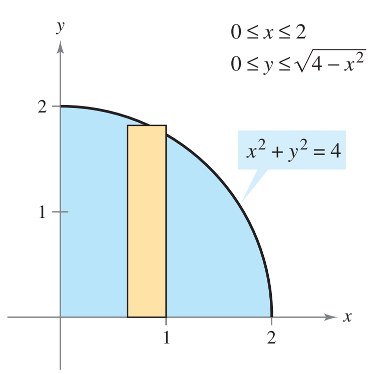


Figure 14.54

Example 2 is unusual in that all six possible orders of integration produce integrals of comparable difficulty. Try setting up some other possible orders of integration to find the volume of the ellipsoid. For instance, the order $dx\,dy\,dz$ yields the integral

$$V = 8\int_0^4 \int_0^{\sqrt{16-z^2}/2} \int_0^{\sqrt{16-4y^2-z^2}/2} dx\,dy\,dz.$$

If you solve this integral, you will obtain the same volume obtained in Example 2. This is always the case—the order of integration does not affect the value of the integral. However, the order of integration often does affect the complexity of the integral. In Example 3, the given order of integration is not convenient, so you can change the order to simplify the problem.

EXAMPLE 3 Changing the Order of Integration

Evaluate $\displaystyle\int_0^{\sqrt{\pi/2}} \int_x^{\sqrt{\pi/2}} \int_1^3 \sin(y^2)\,dz\,dy\,dx.$

Solution Note that after one integration in the given order, you would encounter the integral $2\int \sin(y^2)\,dy$, which is not an elementary function. To avoid this problem, change the order of integration to $dz\,dx\,dy$, so that y is the outer variable. The solid region Q is given by

$$0 \le x \le \sqrt{\frac{\pi}{2}}, \quad x \le y \le \sqrt{\frac{\pi}{2}}, \quad 1 \le z \le 3$$

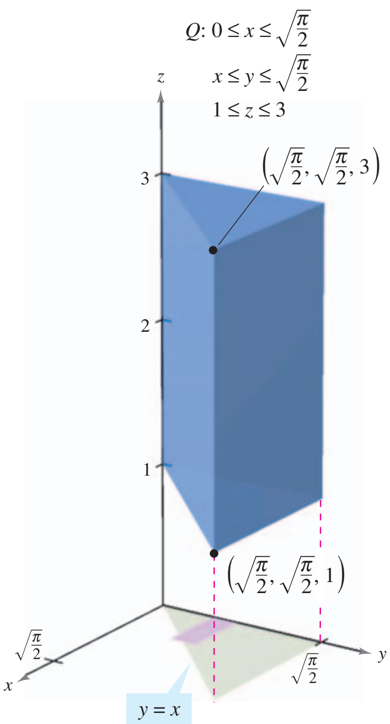
and the projection of Q in the xy-plane yields the bounds

$$0 \le y \le \sqrt{\frac{\pi}{2}} \quad \text{and} \quad 0 \le x \le y.$$

So, you have

$$\begin{aligned}
V &= \iiint_Q dV \\
&= \int_0^{\sqrt{\pi/2}} \int_0^y \int_1^3 \sin(y^2)\,dz\,dx\,dy \\
&= \int_0^{\sqrt{\pi/2}} \int_0^y z\sin(y^2)\Big]_1^3\,dx\,dy \\
&= 2\int_0^{\sqrt{\pi/2}} \int_0^y \sin(y^2)\,dx\,dy \\
&= 2\int_0^{\sqrt{\pi/2}} x\sin(y^2)\Big]_0^y\,dy \\
&= 2\int_0^{\sqrt{\pi/2}} y\sin(y^2)\,dy \\
&= -\cos(y^2)\Big]_0^{\sqrt{\pi/2}} \\
&= 1.
\end{aligned}$$

See Figure 14.55.



The volume of the solid region Q is 1.
Figure 14.55

Rotatable Graph

Try It

Exploration A

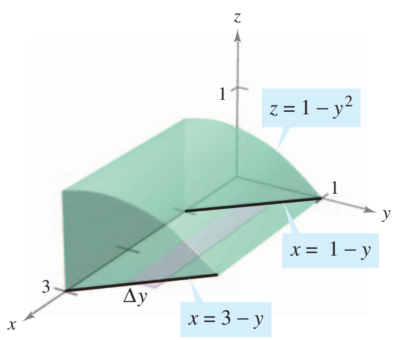
EXAMPLE 4 Determining the Limits of Integration

Set up a triple integral for the volume of each solid region.

a. The region in the first octant bounded above by the cylinder $z = 1 - y^2$ and lying between the vertical planes $x + y = 1$ and $x + y = 3$

b. The upper hemisphere given by $z = \sqrt{1 - x^2 - y^2}$

c. The region bounded below by the paraboloid $z = x^2 + y^2$ and above by the sphere $x^2 + y^2 + z^2 = 6$



Q: $0 \le z \le 1 - y^2$
$1 - y \le x \le 3 - y$
$0 \le y \le 1$

Figure 14.56

Rotatable Graph

Solution

a. In Figure 14.56, note that the solid is bounded below by the xy-plane $(z = 0)$ and above by the cylinder $z = 1 - y^2$. So,

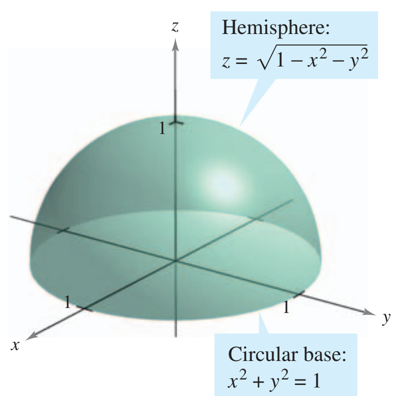
$$0 \le z \le 1 - y^2. \qquad \text{Bounds for } z$$

Projecting the region onto the xy-plane produces a parallelogram. Because two sides of the parallelogram are parallel to the x-axis, you have the following bounds:

$$1 - y \le x \le 3 - y \quad \text{and} \quad 0 \le y \le 1.$$

So, the volume of the region is given by

$$V = \iiint_Q dV = \int_0^1 \int_{1-y}^{3-y} \int_0^{1-y^2} dz\, dx\, dy.$$



Q: $0 \le z \le \sqrt{1 - x^2 - y^2}$
$-\sqrt{1 - y^2} \le x \le \sqrt{1 - y^2}$
$-1 \le y \le 1$

Figure 14.57

Rotatable Graph

b. For the upper hemisphere given by $z = \sqrt{1 - x^2 - y^2}$, you have

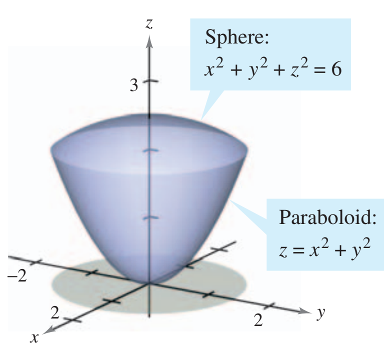
$$0 \le z \le \sqrt{1 - x^2 - y^2}. \qquad \text{Bounds for } z$$

In Figure 14.57, note that the projection of the hemisphere onto the xy-plane is the circle given by $x^2 + y^2 = 1$, and you can use either order $dx\, dy$ or $dy\, dx$. Choosing the first produces

$$-\sqrt{1 - y^2} \le x \le \sqrt{1 - y^2} \quad \text{and} \quad -1 \le y \le 1$$

which implies that the volume of the region is given by

$$V = \iiint_Q dV = \int_{-1}^1 \int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} \int_0^{\sqrt{1-x^2-y^2}} dz\, dx\, dy.$$



Q: $x^2 + y^2 \le z \le \sqrt{6 - x^2 - y^2}$
$-\sqrt{2 - x^2} \le y \le \sqrt{2 - x^2}$
$-\sqrt{2} \le x \le \sqrt{2}$

Figure 14.58

Rotatable Graph

c. For the region bounded below by the paraboloid $z = x^2 + y^2$ and above by the sphere $x^2 + y^2 + z^2 = 6$, you have

$$x^2 + y^2 \le z \le \sqrt{6 - x^2 - y^2}. \qquad \text{Bounds for } z$$

The sphere and the paraboloid intersect when $z = 2$. Moreover, you can see in Figure 14.58 that the projection of the solid region onto the xy-plane is the circle given by $x^2 + y^2 = 2$. Using the order $dy\, dx$ produces

$$-\sqrt{2 - x^2} \le y \le \sqrt{2 - x^2} \quad \text{and} \quad -\sqrt{2} \le x \le \sqrt{2}$$

which implies that the volume of the region is given by

$$V = \iiint_Q dV = \int_{-\sqrt{2}}^{\sqrt{2}} \int_{-\sqrt{2-x^2}}^{\sqrt{2-x^2}} \int_{x^2+y^2}^{\sqrt{6-x^2-y^2}} dz\, dy\, dx.$$

Try It

Exploration A

Center of Mass and Moments of Inertia

EXPLORATION

Sketch the solid (of uniform density) bounded by $z = 0$ and

$$z = \frac{1}{1 + x^2 + y^2}$$

where $x^2 + y^2 \leq 1$. From your sketch, estimate the coordinates of the center of mass of the solid. Now use a computer algebra system to verify your estimate. What do you observe?

In the remainder of this section, two important engineering applications of triple integrals are discussed. Consider a solid region Q whose density is given by the **density function ρ.** The **center of mass** of a solid region Q of mass m is given by $(\bar{x}, \bar{y}, \bar{z})$, where

$$m = \iiint_Q \rho(x, y, z)\, dV \qquad \text{Mass of the solid}$$

$$M_{yz} = \iiint_Q x\rho(x, y, z)\, dV \qquad \text{First moment about } yz\text{-plane}$$

$$M_{xz} = \iiint_Q y\rho(x, y, z)\, dV \qquad \text{First moment about } xz\text{-plane}$$

$$M_{xy} = \iiint_Q z\rho(x, y, z)\, dV \qquad \text{First moment about } xy\text{-plane}$$

and

$$\bar{x} = \frac{M_{yz}}{m}, \quad \bar{y} = \frac{M_{xz}}{m}, \quad \bar{z} = \frac{M_{xy}}{m}.$$

The quantities M_{yz}, M_{xz}, and M_{xy} are called the **first moments** of the region Q about the yz-, xz-, and xy-planes, respectively.

The first moments for solid regions are taken about a plane, whereas the second moments for solids are taken about a line. The **second moments** (or **moments of inertia**) about the x-, y-, and z-axes are as follows.

NOTE In engineering and physics, the moment of inertia of a mass is used to find the time required for a mass to reach a given speed of rotation about an axis, as shown in Figure 14.59. The greater the moment of inertia, the longer a force must be applied for the mass to reach the given speed.

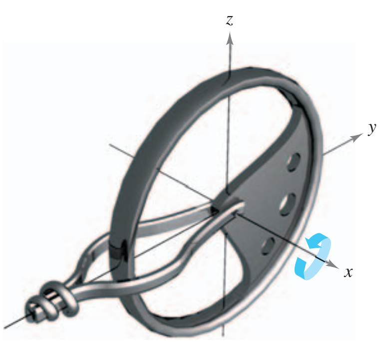


Figure 14.59

Rotatable Graph

$$I_x = \iiint_Q (y^2 + z^2)\rho(x, y, z)\, dV \qquad \text{Moment of inertia about } x\text{-axis}$$

$$I_y = \iiint_Q (x^2 + z^2)\rho(x, y, z)\, dV \qquad \text{Moment of inertia about } y\text{-axis}$$

$$I_z = \iiint_Q (x^2 + y^2)\rho(x, y, z)\, dV \qquad \text{Moment of inertia about } z\text{-axis}$$

For problems requiring the calculation of all three moments, considerable effort can be saved by applying the additive property of triple integrals and writing

$$I_x = I_{xz} + I_{xy}, \quad I_y = I_{yz} + I_{xy}, \quad \text{and} \quad I_z = I_{yz} + I_{xz}$$

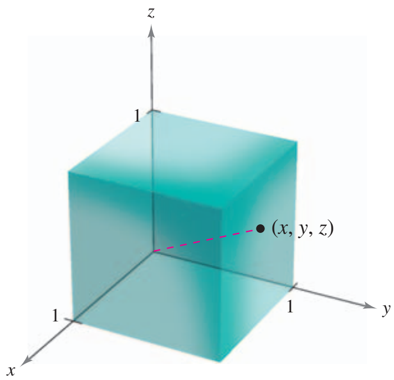
where I_{xy}, I_{xz}, and I_{yz} are as follows.

$$I_{xy} = \iiint_Q z^2\rho(x, y, z)\, dV$$

$$I_{xz} = \iiint_Q y^2\rho(x, y, z)\, dV$$

$$I_{yz} = \iiint_Q x^2\rho(x, y, z)\, dV$$

EXAMPLE 5 Finding the Center of Mass of a Solid Region



Variable density:
$\rho(x, y, z) = k(x^2 + y^2 + z^2)$
Figure 14.60

Rotatable Graph

Find the center of mass of the unit cube shown in Figure 14.60, given that the density at the point (x, y, z) is proportional to the square of its distance from the origin.

Solution Because the density at (x, y, z) is proportional to the square of the distance between $(0, 0, 0)$ and (x, y, z), you have

$$\rho(x, y, z) = k(x^2 + y^2 + z^2).$$

You can use this density function to find the mass of the cube. Because of the symmetry of the region, any order of integration will produce an integral of comparable difficulty.

$$\begin{aligned} m &= \int_0^1\int_0^1\int_0^1 k(x^2 + y^2 + z^2)\,dz\,dy\,dx \\ &= k\int_0^1\int_0^1 \left[(x^2 + y^2)z + \frac{z^3}{3}\right]_0^1 dy\,dx \\ &= k\int_0^1\int_0^1 \left(x^2 + y^2 + \frac{1}{3}\right) dy\,dx \\ &= k\int_0^1 \left[\left(x^2 + \frac{1}{3}\right)y + \frac{y^3}{3}\right]_0^1 dx \\ &= k\int_0^1 \left(x^2 + \frac{2}{3}\right) dx \\ &= k\left[\frac{x^3}{3} + \frac{2x}{3}\right]_0^1 = k \end{aligned}$$

The first moment about the yz-plane is

$$\begin{aligned} M_{yz} &= k\int_0^1\int_0^1\int_0^1 x(x^2 + y^2 + z^2)\,dz\,dy\,dx \\ &= k\int_0^1 x\left[\int_0^1\int_0^1 (x^2 + y^2 + z^2)\,dz\,dy\right] dx. \end{aligned}$$

Note that x can be factored out of the two inner integrals, because it is constant with respect to y and z. After factoring, the two inner integrals are the same as for the mass m. Therefore, you have

$$\begin{aligned} M_{yz} &= k\int_0^1 x\left(x^2 + \frac{2}{3}\right) dx \\ &= k\left[\frac{x^4}{4} + \frac{x^2}{3}\right]_0^1 \\ &= \frac{7k}{12}. \end{aligned}$$

So,

$$\bar{x} = \frac{M_{yz}}{m} = \frac{7k/12}{k} = \frac{7}{12}.$$

Finally, from the nature of ρ and the symmetry of x, y, and z in this solid region, you have $\bar{x} = \bar{y} = \bar{z}$, and the center of mass is $\left(\frac{7}{12}, \frac{7}{12}, \frac{7}{12}\right)$.

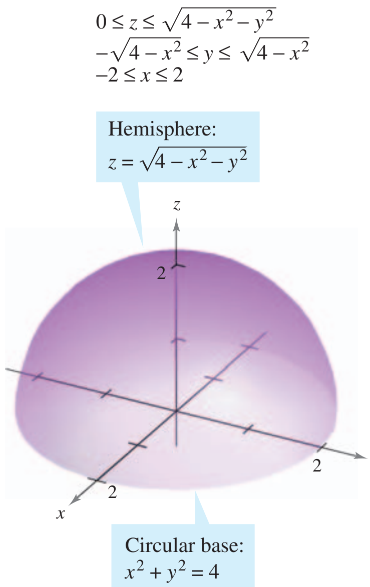
Try It Open Exploration

EXAMPLE 6 Moments of Inertia for a Solid Region

Find the moments of inertia about the x- and y-axes for the solid region lying between the hemisphere

$$z = \sqrt{4 - x^2 - y^2}$$

and the xy-plane, given that the density at (x, y, z) is proportional to the distance between (x, y, z) and the xy-plane.



Variable density: $\rho(x, y, z) = kz$
Figure 14.61

Rotatable Graph

Solution The density of the region is given by $\rho(x, y, z) = kz$. Considering the symmetry of this problem, you know that $I_x = I_y$, and you need to compute only one moment, say I_x. From Figure 14.61, choose the order $dz\,dy\,dx$ and write

$$\begin{aligned}
I_x &= \iiint\limits_Q (y^2 + z^2)\rho(x, y, z)\,dV \\
&= \int_{-2}^{2}\int_{-\sqrt{4-x^2}}^{\sqrt{4-x^2}}\int_{0}^{\sqrt{4-x^2-y^2}} (y^2 + z^2)(kz)\,dz\,dy\,dx \\
&= k\int_{-2}^{2}\int_{-\sqrt{4-x^2}}^{\sqrt{4-x^2}} \left[\frac{y^2z^2}{2} + \frac{z^4}{4}\right]_0^{\sqrt{4-x^2-y^2}} dy\,dx \\
&= k\int_{-2}^{2}\int_{-\sqrt{4-x^2}}^{\sqrt{4-x^2}} \left[\frac{y^2(4 - x^2 - y^2)}{2} + \frac{(4 - x^2 - y^2)^2}{4}\right] dy\,dx \\
&= \frac{k}{4}\int_{-2}^{2}\int_{-\sqrt{4-x^2}}^{\sqrt{4-x^2}} [(4 - x^2)^2 - y^4]\,dy\,dx \\
&= \frac{k}{4}\int_{-2}^{2} \left[(4 - x^2)^2 y - \frac{y^5}{5}\right]_{-\sqrt{4-x^2}}^{\sqrt{4-x^2}} dx \\
&= \frac{k}{4}\int_{-2}^{2} \frac{8}{5}(4 - x^2)^{5/2}\,dx \\
&= \frac{4k}{5}\int_{0}^{2} (4 - x^2)^{5/2}\,dx && x = 2\sin\theta \\
&= \frac{4k}{5}\int_{0}^{\pi/2} 64\cos^6\theta\,d\theta \\
&= \left(\frac{256k}{5}\right)\left(\frac{5\pi}{32}\right) && \text{Wallis's Formula} \\
&= 8k\pi.
\end{aligned}$$

So, $I_x = 8k\pi = I_y$.

Exploration A

In Example 6, notice that the moments of inertia about the x- and y-axes are equal to each other. The moment about the z-axis, however, is different. Does it seem that the moment of inertia about the z-axis should be less than or greater than the moments calculated in Example 6? By performing the calculations, you can determine that

$$I_z = \frac{16}{3}k\pi.$$

This tells you that the solid shown in Figure 14.61 has a greater resistance to rotation about the x- or y-axis than about the z-axis.

Section 14.7 Triple Integrals in Cylindrical and Spherical Coordinates

- **Write and evaluate a triple integral in cylindrical coordinates.**
- **Write and evaluate a triple integral in spherical coordinates.**

PIERRE SIMON DE LAPLACE (1749–1827)

One of the first to use a cylindrical coordinate system was the French mathematician Pierre Simon de Laplace. Laplace has been called the "Newton of France," and he published many important works in mechanics, differential equations, and probability.

MathBio

Triple Integrals in Cylindrical Coordinates

Many common solid regions such as spheres, ellipsoids, cones, and paraboloids can yield difficult triple integrals in rectangular coordinates. In fact, it is precisely this difficulty that led to the introduction of nonrectangular coordinate systems. In this section, you will learn how to use *cylindrical* and *spherical* coordinates to evaluate triple integrals.

Recall from Section 11.7 that the rectangular conversion equations for cylindrical coordinates are

$$x = r\cos\theta$$
$$y = r\sin\theta$$
$$z = z.$$

STUDY TIP An easy way to remember these conversions is to note that the equations for x and y are the same as in polar coordinates and z is unchanged.

In this coordinate system, the simplest solid region is a cylindrical block determined by

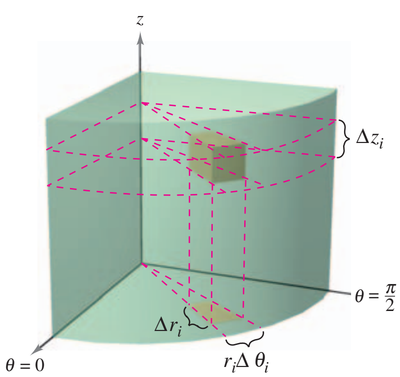
$$r_1 \le r \le r_2, \quad \theta_1 \le \theta \le \theta_2, \quad z_1 \le z \le z_2$$

as shown in Figure 14.62. To obtain the cylindrical coordinate form of a triple integral, suppose that Q is a solid region whose projection R onto the xy-plane can be described in polar coordinates. That is,

$$Q = \{(x, y, z): (x, y) \text{ is in } R, \quad h_1(x, y) \le z \le h_2(x, y)\}$$

and

$$R = \{(r, \theta): \theta_1 \le \theta \le \theta_2, \quad g_1(\theta) \le r \le g_2(\theta)\}.$$



Volume of cylindrical block:
$\Delta V_i = r_i \Delta r_i \Delta\theta_i \Delta z_i$
Figure 14.62

Rotatable Graph

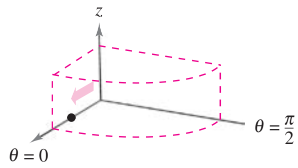
If f is a continuous function on the solid Q, you can write the triple integral of f over Q as

$$\iiint_Q f(x, y, z)\, dV = \iint_R \left[\int_{h_1(x, y)}^{h_2(x, y)} f(x, y, z)\, dz\right] dA$$

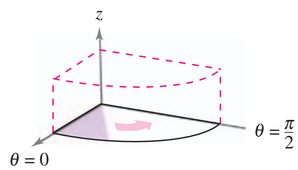
where the double integral over R is evaluated in polar coordinates. That is, R is a plane region that is either r-simple or θ-simple. If R is r-simple, the iterated form of the triple integral in cylindrical form is

$$\iiint_Q f(x, y, z)\, dV = \int_{\theta_1}^{\theta_2}\int_{g_1(\theta)}^{g_2(\theta)}\int_{h_1(r\cos\theta,\, r\sin\theta)}^{h_2(r\cos\theta,\, r\sin\theta)} f(r\cos\theta, r\sin\theta, z)r\, dz\, dr\, d\theta.$$

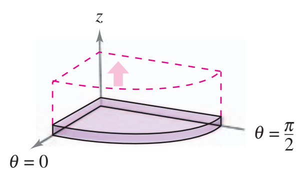
NOTE This is only one of six possible orders of integration. The other five are $dz\, d\theta\, dr$, $dr\, dz\, d\theta$, $dr\, d\theta\, dz$, $d\theta\, dz\, dr$, and $d\theta\, dr\, dz$.



Integrate with respect to r.



Integrate with respect to θ.



Integrate with respect to z.

Figure 14.63

Animation

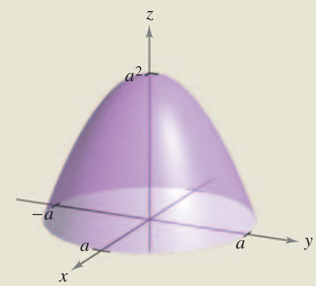
To visualize a particular order of integration, it helps to view the iterated integral in terms of three sweeping motions—each adding another dimension to the solid. For instance, in the order $dr\,d\theta\,dz$, the first integration occurs in the r-direction as a point sweeps out a ray. Then, as θ increases, the line sweeps out a sector. Finally, as z increases, the sector sweeps out a solid wedge, as shown in Figure 14.63.

EXPLORATION

Volume of a Paraboloid Sector On pages 995, 1003, and 1025, you were asked to summarize the different ways you know for finding the volume of the solid bounded by the paraboloid

$$z = a^2 - x^2 - y^2, \quad a > 0$$

and the xy-plane. You now know one more way. Use it to find the volume of the solid. Compare the different methods. What are the advantages and disadvantages of each?



Rotatable Graph

EXAMPLE 1 Finding Volume by Cylindrical Coordinates

Find the volume of the solid region Q cut from the sphere

$$x^2 + y^2 + z^2 = 4 \qquad \text{Sphere}$$

by the cylinder $r = 2 \sin \theta$, as shown in Figure 14.64.

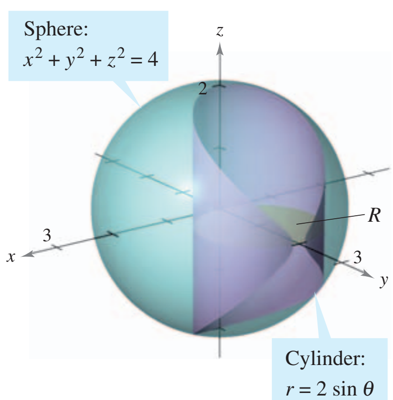


Figure 14.64

Rotatable Graph

Solution Because $x^2 + y^2 + z^2 = r^2 + z^2 = 4$, the bounds on z are

$$-\sqrt{4 - r^2} \le z \le \sqrt{4 - r^2}.$$

Let R be the circular projection of the solid onto the $r\theta$-plane. Then the bounds on R are $0 \le r \le 2 \sin \theta$ and $0 \le \theta \le \pi$. So, the volume of Q is

$$\begin{aligned}
V &= \int_0^{\pi} \int_0^{2\sin\theta} \int_{-\sqrt{4-r^2}}^{\sqrt{4-r^2}} r\,dz\,dr\,d\theta \\
&= 2\int_0^{\pi/2} \int_0^{2\sin\theta} 2r\sqrt{4 - r^2}\,dr\,d\theta \\
&= 2\int_0^{\pi/2} -\frac{2}{3}(4 - r^2)^{3/2}\Big]_0^{2\sin\theta}\,d\theta \\
&= \frac{4}{3}\int_0^{\pi/2} (8 - 8\cos^3\theta)\,d\theta \\
&= \frac{32}{3}\int_0^{\pi/2} [1 - (\cos\theta)(1 - \sin^2\theta)]\,d\theta \\
&= \frac{32}{3}\left[\theta - \sin\theta + \frac{\sin^3\theta}{3}\right]_0^{\pi/2} \\
&= \frac{16}{9}(3\pi - 4) \\
&\approx 9.644.
\end{aligned}$$

Try It   Exploration A   Exploration B

EXAMPLE 2 Finding Mass by Cylindrical Coordinates

Find the mass of the ellipsoid Q given by $4x^2 + 4y^2 + z^2 = 16$, lying above the xy-plane. The density at a point in the solid is proportional to the distance between the point and the xy-plane.

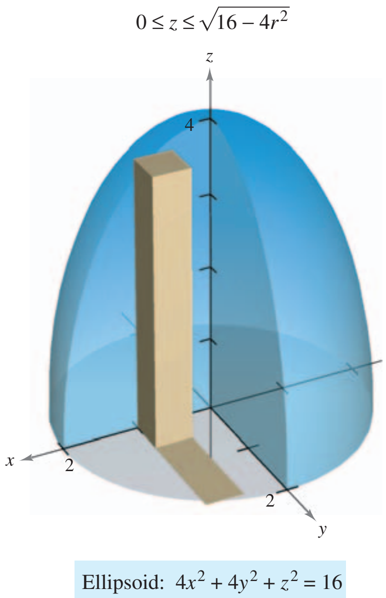


Figure 14.65

Rotatable Graph

Solution The density function is $\rho(r, \theta, z) = kz$. The bounds on z are

$$0 \le z \le \sqrt{16 - 4x^2 - 4y^2} = \sqrt{16 - 4r^2}$$

where $0 \le r \le 2$ and $0 \le \theta \le 2\pi$, as shown in Figure 14.65. The mass of the solid is

$$\begin{aligned} m &= \int_0^{2\pi}\int_0^2\int_0^{\sqrt{16-4r^2}} kzr\,dz\,dr\,d\theta \\ &= \frac{k}{2}\int_0^{2\pi}\int_0^2 z^2 r\Big]_0^{\sqrt{16-4r^2}}\,dr\,d\theta \\ &= \frac{k}{2}\int_0^{2\pi}\int_0^2 (16r - 4r^3)\,dr\,d\theta \\ &= \frac{k}{2}\int_0^{2\pi}\Big[8r^2 - r^4\Big]_0^2\,d\theta \\ &= 8k\int_0^{2\pi} d\theta = 16\pi k. \end{aligned}$$

Try It Exploration A

Integration in cylindrical coordinates is useful when factors involving $x^2 + y^2$ appear in the integrand, as illustrated in Example 3.

EXAMPLE 3 Finding a Moment of Inertia

Find the moment of inertia about the axis of symmetry of the solid Q bounded by the paraboloid $z = x^2 + y^2$ and the plane $z = 4$, as shown in Figure 14.66. The density at each point is proportional to the distance between the point and the z-axis.

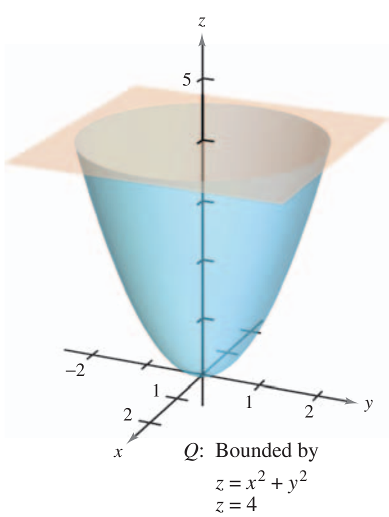


Figure 14.66

Rotatable Graph

Solution Because the z-axis is the axis of symmetry, and $\rho(x, y, z) = k\sqrt{x^2 + y^2}$, it follows that

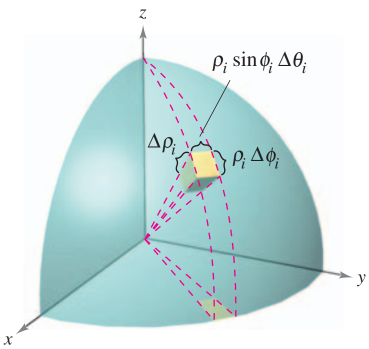
$$I_z = \iiint_Q k(x^2 + y^2)\sqrt{x^2 + y^2}\,dV.$$

In cylindrical coordinates, $0 \le r \le \sqrt{x^2 + y^2} = \sqrt{z}$. So, you have

$$\begin{aligned} I_z &= k\int_0^4\int_0^{2\pi}\int_0^{\sqrt{z}} r^2(r)r\,dr\,d\theta\,dz \\ &= k\int_0^4\int_0^{2\pi}\frac{r^5}{5}\Big]_0^{\sqrt{z}}\,d\theta\,dz \\ &= k\int_0^4\int_0^{2\pi}\frac{z^{5/2}}{5}\,d\theta\,dz \\ &= \frac{k}{5}\int_0^4 z^{5/2}(2\pi)\,dz \\ &= \frac{2\pi k}{5}\left[\frac{2}{7}z^{7/2}\right]_0^4 = \frac{512k\pi}{35}. \end{aligned}$$

Try It Exploration A

Triple Integrals in Spherical Coordinates



Spherical block:
$\Delta V_i \approx \rho_i^2 \sin\phi_i \,\Delta\rho_i \,\Delta\theta_i \,\Delta\phi_i$
Figure 14.67

Rotatable Graph

Triple integrals involving spheres or cones are often easier to evaluate by converting to spherical coordinates. Recall from Section 11.7 that the rectangular conversion equations for spherical coordinates are

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

In this coordinate system, the simplest region is a spherical block determined by

$$\{(\rho, \theta, \phi): \rho_1 \le \rho \le \rho_2, \quad \theta_1 \le \theta \le \theta_2, \quad \phi_1 \le \phi \le \phi_2\}$$

where $\rho_1 \ge 0$, $\theta_2 - \theta_1 \le 2\pi$, and $0 \le \phi_1 \le \phi_2 \le \pi$, as shown in Figure 14.67. If (ρ, θ, ϕ) is a point in the interior of such a block, then the volume of the block can be approximated by $\Delta V \approx \rho^2 \sin\phi \,\Delta\rho \,\Delta\phi \,\Delta\theta$ (see Exercise 17 in the Problem Solving exercises for this chapter).

Using the usual process involving an inner partition, summation, and a limit, you can develop the following version of a triple integral in spherical coordinates for a continuous function f defined on the solid region Q.

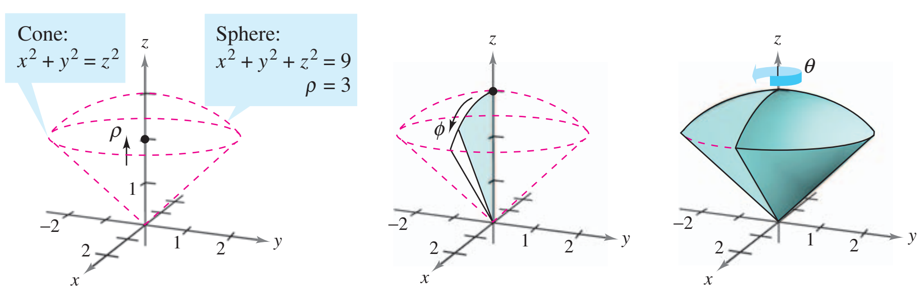
$$\iiint_Q f(x, y, z)\, dV = \int_{\theta_1}^{\theta_2}\int_{\phi_1}^{\phi_2}\int_{\rho_1}^{\rho_2} f(\rho\sin\phi\cos\theta, \rho\sin\phi\sin\theta, \rho\cos\phi)\rho^2 \sin\phi \, d\rho\, d\phi\, d\theta.$$

This formula can be modified for different orders of integration and generalized to include regions with variable boundaries.

Like triple integrals in cylindrical coordinates, triple integrals in spherical coordinates are evaluated with iterated integrals. As with cylindrical coordinates, you can visualize a particular order of integration by viewing the iterated integral in terms of three sweeping motions—each adding another dimension to the solid. For instance, the iterated integral

$$\int_0^{2\pi}\int_0^{\pi/4}\int_0^{3} \rho^2 \sin\phi \, d\rho\, d\phi\, d\theta$$

(which is used in Example 4) is illustrated in Figure 14.68.



ρ varies from 0 to 3 with ϕ and θ held constant.
ϕ varies from 0 to $\pi/4$ with θ held constant.
θ varies from 0 to 2π.
Figure 14.68

Animation

NOTE The Greek letter ρ used in spherical coordinates is not related to density. Rather, it is the three-dimensional analog of the r used in polar coordinates. For problems involving spherical coordinates and a density function, this text uses a different symbol to denote density.

EXAMPLE 4 Finding Volume in Spherical Coordinates

Find the volume of the solid region Q bounded below by the upper nappe of the cone $z^2 = x^2 + y^2$ and above by the sphere $x^2 + y^2 + z^2 = 9$, as shown in Figure 14.69.

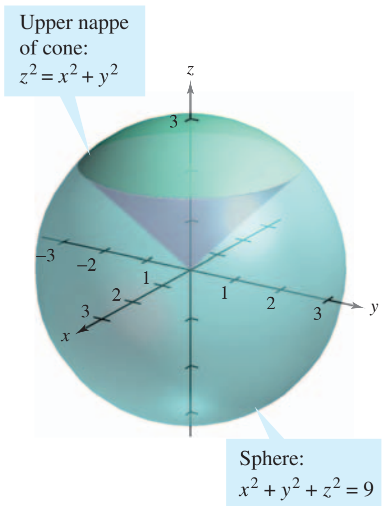


Figure 14.69

Rotatable Graph

Solution In spherical coordinates, the equation of the sphere is

$$\rho^2 = x^2 + y^2 + z^2 = 9 \quad \Rightarrow \quad \rho = 3.$$

Furthermore, the sphere and cone intersect when

$$(x^2 + y^2) + z^2 = (z^2) + z^2 = 9 \quad \Rightarrow \quad z = \frac{3}{\sqrt{2}}$$

and, because $z = \rho \cos \phi$, it follows that

$$\left(\frac{3}{\sqrt{2}}\right)\left(\frac{1}{3}\right) = \cos \phi \quad \Rightarrow \quad \phi = \frac{\pi}{4}.$$

Consequently, you can use the integration order $d\rho\, d\phi\, d\theta$, where $0 \le \rho \le 3$, $0 \le \phi \le \pi/4$, and $0 \le \theta \le 2\pi$. The volume is

$$\begin{aligned} V = \iiint_Q dV &= \int_0^{2\pi} \int_0^{\pi/4} \int_0^3 \rho^2 \sin \phi \, d\rho\, d\phi\, d\theta \\ &= \int_0^{2\pi} \int_0^{\pi/4} 9 \sin \phi \, d\phi\, d\theta \\ &= 9 \int_0^{2\pi} -\cos \phi \Big]_0^{\pi/4} d\theta \\ &= 9 \int_0^{2\pi} \left(1 - \frac{\sqrt{2}}{2}\right) d\theta = 9\pi\left(2 - \sqrt{2}\right) \approx 16.563. \end{aligned}$$

Try It Exploration A

EXAMPLE 5 Finding the Center of Mass of a Solid Region

Find the center of mass of the solid region Q of uniform density, bounded below by the upper nappe of the cone $z^2 = x^2 + y^2$ and above by the sphere $x^2 + y^2 + z^2 = 9$.

Solution Because the density is uniform, you can consider the density at the point (x, y, z) to be k. By symmetry, the center of mass lies on the z-axis, and you need only calculate $\bar{z} = M_{xy}/m$, where $m = kV = 9k\pi\left(2 - \sqrt{2}\right)$ from Example 4. Because $z = \rho \cos \phi$, it follows that

$$\begin{aligned} M_{xy} = \iiint_Q kz\, dV &= k \int_0^3 \int_0^{2\pi} \int_0^{\pi/4} (\rho \cos \phi)\rho^2 \sin \phi \, d\phi\, d\theta\, d\rho \\ &= k \int_0^3 \int_0^{2\pi} \rho^3 \frac{\sin^2 \phi}{2}\Big]_0^{\pi/4} d\theta\, d\rho \\ &= \frac{k}{4} \int_0^3 \int_0^{2\pi} \rho^3 \, d\theta\, d\rho = \frac{k\pi}{2} \int_0^3 \rho^3 \, d\rho = \frac{81k\pi}{8}. \end{aligned}$$

So,

$$\bar{z} = \frac{M_{xy}}{m} = \frac{81k\pi/8}{9k\pi\left(2 - \sqrt{2}\right)} = \frac{9\left(2 + \sqrt{2}\right)}{16} \approx 1.920$$

and the center of mass is approximately $(0, 0, 1.92)$.

Try It Exploration A Open Exploration

Section 14.8 Change of Variables: Jacobians

- **Understand the concept of a Jacobian.**
- **Use a Jacobian to change variables in a double integral.**

Carl Gustav Jacobi (1804–1851)

The Jacobian is named after the German mathematician Carl Gustav Jacobi. Jacobi is known for his work in many areas of mathematics, but his interest in integration stemmed from the problem of finding the circumference of an ellipse.

MathBio

Jacobians

For the single integral

$$\int_a^b f(x)\,dx$$

you can change variables by letting $x = g(u)$, so that $dx = g'(u)\,du$, and obtain

$$\int_a^b f(x)\,dx = \int_c^d f(g(u))g'(u)\,du$$

where $a = g(c)$ and $b = g(d)$. Note that the change-of-variables process introduces an additional factor $g'(u)$ into the integrand. This also occurs in the case of double integrals

$$\int_R\int f(x, y)\,dA = \int_S\int f(g(u, v), h(u, v))\underbrace{\left|\frac{\partial x}{\partial u}\frac{\partial y}{\partial v} - \frac{\partial y}{\partial u}\frac{\partial x}{\partial v}\right|}_{\text{Jacobian}}du\,dv$$

where the change of variables $x = g(u, v)$ and $y = h(u, v)$ introduces a factor called the **Jacobian** of x and y with respect to u and v. In defining the Jacobian, it is convenient to use the following determinant notation.

Definition of the Jacobian

If $x = g(u, v)$ and $y = h(u, v)$, then the **Jacobian** of x and y with respect to u and v, denoted by $\partial(x, y)/\partial(u, v)$, is

$$\frac{\partial(x, y)}{\partial(u, v)} = \begin{vmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial x}{\partial v} \\ \dfrac{\partial y}{\partial u} & \dfrac{\partial y}{\partial v} \end{vmatrix} = \frac{\partial x}{\partial u}\frac{\partial y}{\partial v} - \frac{\partial y}{\partial u}\frac{\partial x}{\partial v}.$$

EXAMPLE 1 The Jacobian for Rectangular-to-Polar Conversion

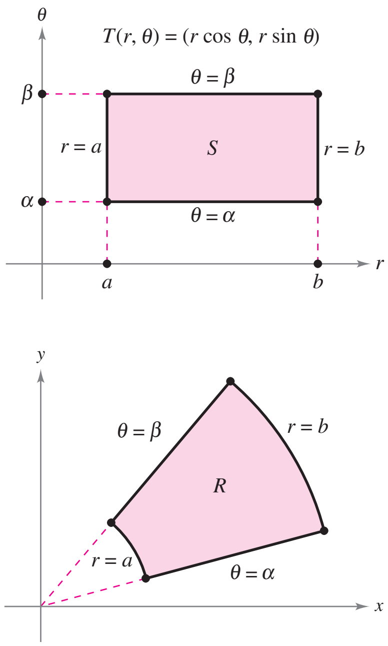
Find the Jacobian for the change of variables defined by

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

Solution From the definition of a Jacobian, you obtain

$$\begin{aligned}\frac{\partial(x, y)}{\partial(r, \theta)} &= \begin{vmatrix} \dfrac{\partial x}{\partial r} & \dfrac{\partial x}{\partial \theta} \\ \dfrac{\partial y}{\partial r} & \dfrac{\partial y}{\partial \theta} \end{vmatrix} \\ &= \begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix} \\ &= r\cos^2\theta + r\sin^2\theta \\ &= r.\end{aligned}$$

Try It Exploration A Exploration B



S is the region in the *rθ*-plane that corresponds to *R* in the *xy*-plane.
Figure 14.70

Example 1 points out that the change of variables from rectangular to polar coordinates for a double integral can be written as

$$\int_R\!\!\int f(x, y)\, dA = \int_S\!\!\int f(r\cos\theta, r\sin\theta) r\, dr\, d\theta, \quad r > 0$$
$$= \int_S\!\!\int f(r\cos\theta, r\sin\theta)\left|\frac{\partial(x, y)}{\partial(r, \theta)}\right| dr\, d\theta$$

where S is the region in the $r\theta$-plane that corresponds to the region R in the xy-plane, as shown in Figure 14.70. This formula is similar to that found on page 1003.

In general, a change of variables is given by a one-to-one **transformation** T from a region S in the uv-plane to a region R in the xy-plane, to be given by

$$T(u, v) = (x, y) = (g(u, v), h(u, v))$$

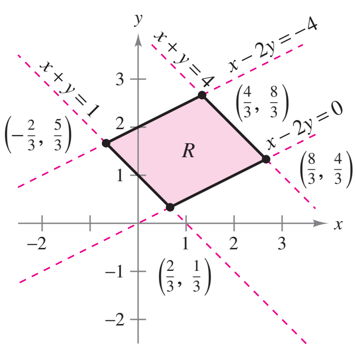
where g and h have continuous first partial derivatives in the region S. Note that the point (u, v) lies in S and the point (x, y) lies in R. In most cases, you are hunting for a transformation in which the region S is simpler than the region R.

EXAMPLE 2 Finding a Change of Variables to Simplify a Region

Let R be the region bounded by the lines

$$x - 2y = 0, \qquad x - 2y = -4, \qquad x + y = 4, \qquad \text{and} \qquad x + y = 1$$

as shown in Figure 14.71. Find a transformation T from a region S to R such that S is a rectangular region (with sides parallel to the u- or v-axis).



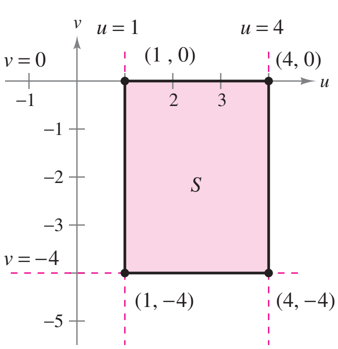
Region R in the xy-plane
Figure 14.71

Solution To begin, let $u = x + y$ and $v = x - 2y$. Solving this system of equations for x and y produces $T(u, v) = (x, y)$, where

$$x = \frac{1}{3}(2u + v) \qquad \text{and} \qquad y = \frac{1}{3}(u - v).$$

The four boundaries for R in the xy-plane give rise to the following bounds for S in the uv-plane.

Bounds in the xy-Plane		*Bounds in the uv-Plane*
$x + y = 1$	⇨	$u = 1$
$x + y = 4$	⇨	$u = 4$
$x - 2y = 0$	⇨	$v = 0$
$x - 2y = -4$	⇨	$v = -4$



Region S in the uv-plane
Figure 14.72

The region S is shown in Figure 14.72. Note that the transformation T maps the vertices of the region S onto the vertices of the region R. For instance,

$$T(1, 0) = \left(\tfrac{1}{3}[2(1) + 0], \tfrac{1}{3}[1 - 0]\right) = \left(\tfrac{2}{3}, \tfrac{1}{3}\right)$$
$$T(4, 0) = \left(\tfrac{1}{3}[2(4) + 0], \tfrac{1}{3}[4 - 0]\right) = \left(\tfrac{8}{3}, \tfrac{4}{3}\right)$$
$$T(4, -4) = \left(\tfrac{1}{3}[2(4) - 4], \tfrac{1}{3}[4 - (-4)]\right) = \left(\tfrac{4}{3}, \tfrac{8}{3}\right)$$
$$T(1, -4) = \left(\tfrac{1}{3}[2(1) - 4], \tfrac{1}{3}[1 - (-4)]\right) = \left(-\tfrac{2}{3}, \tfrac{5}{3}\right).$$

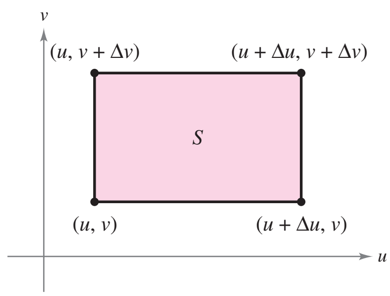
Try It **Exploration A**

Change of Variables for Double Integrals

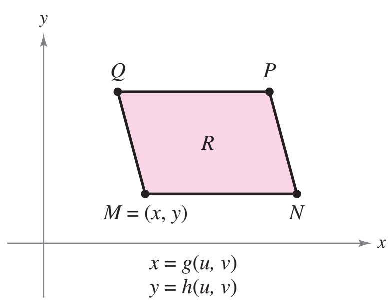
THEOREM 14.5 Change of Variables for Double Integrals

Let R and S be regions in the xy- and uv-planes that are related by the equations $x = g(u, v)$ and $y = h(u, v)$ such that each point in R is the image of a unique point in S. If f is continuous on R, g and h have continuous partial derivatives on S, and $\partial(x, y)/\partial(u, v)$ is nonzero on S, then

$$\int_R\int f(x, y)\,dx\,dy = \int_S\int f(g(u, v), h(u, v))\left|\frac{\partial(x, y)}{\partial(u, v)}\right| du\,dv.$$



Area of $S = \Delta u\,\Delta v$
$\Delta u > 0$, $\Delta v > 0$
Figure 14.73



The vertices in the xy-plane are $M(g(u, v), h(u, v))$, $N(g(u + \Delta u, v), h(u + \Delta u, v))$, $P(g(u + \Delta u, v + \Delta v), h(u + \Delta u, v + \Delta v))$, and $Q(g(u, v + \Delta v), h(u, v + \Delta v))$.
Figure 14.74

Proof Consider the case in which S is a rectangular region in the uv-plane with vertices (u, v), $(u + \Delta u, v)$, $(u + \Delta u, v + \Delta v)$, and $(u, v + \Delta v)$, as shown in Figure 14.73. The images of these vertices in the xy-plane are shown in Figure 14.74. If Δu and Δv are small, the continuity of g and h implies that R is approximately a parallelogram determined by the vectors $\overrightarrow{MN}$ and $\overrightarrow{MQ}$. So, the area of R is

$$\Delta A \approx \|\overrightarrow{MN} \times \overrightarrow{MQ}\|.$$

Moreover, for small Δu and Δv, the partial derivatives of g and h with respect to u can be approximated by

$$g_u(u, v) \approx \frac{g(u + \Delta u, v) - g(u, v)}{\Delta u}$$

and

$$h_u(u, v) \approx \frac{h(u + \Delta u, v) - h(u, v)}{\Delta u}.$$

Consequently,

$$\begin{aligned}\overrightarrow{MN} &= [g(u + \Delta u, v) - g(u, v)]\mathbf{i} + [h(u + \Delta u, v) - h(u, v)]\mathbf{j}\\ &\approx [g_u(u, v)\,\Delta u]\mathbf{i} + [h_u(u, v)\,\Delta u]\mathbf{j}\\ &= \frac{\partial x}{\partial u}\Delta u\mathbf{i} + \frac{\partial y}{\partial u}\Delta u\mathbf{j}.\end{aligned}$$

Similarly, you can approximate $\overrightarrow{MQ}$ by $\dfrac{\partial x}{\partial v}\Delta v\mathbf{i} + \dfrac{\partial y}{\partial v}\Delta v\mathbf{j}$, which implies that

$$\overrightarrow{MN} \times \overrightarrow{MQ} \approx \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \dfrac{\partial x}{\partial u}\Delta u & \dfrac{\partial y}{\partial u}\Delta u & 0 \\ \dfrac{\partial x}{\partial v}\Delta v & \dfrac{\partial y}{\partial v}\Delta v & 0 \end{vmatrix} = \begin{vmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial y}{\partial u} \\ \dfrac{\partial x}{\partial v} & \dfrac{\partial y}{\partial v} \end{vmatrix}\Delta u\,\Delta v\mathbf{k}.$$

It follows that, in Jacobian notation,

$$\Delta A \approx \|\overrightarrow{MN} \times \overrightarrow{MQ}\| \approx \left|\frac{\partial(x, y)}{\partial(u, v)}\right|\Delta u\,\Delta v.$$

Because this approximation improves as Δu and Δv approach 0, the limiting case can be written as

$$dA \approx \|\overrightarrow{MN} \times \overrightarrow{MQ}\| \approx \left|\frac{\partial(x, y)}{\partial(u, v)}\right| du\,dv.$$

The next two examples show how a change of variables can simplify the integration process. The simplification can occur in various ways. You can make a change of variables to simplify either the *region* R or the *integrand* $f(x, y)$, or both.

EXAMPLE 3 Using a Change of Variables to Simplify a Region

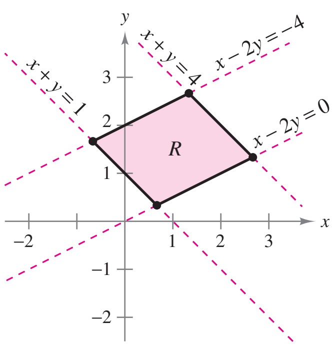


Figure 14.75

Let R be the region bounded by the lines

$$x - 2y = 0, \quad x - 2y = -4, \quad x + y = 4, \quad \text{and} \quad x + y = 1$$

as shown in Figure 14.75. Evaluate the double integral

$$\int_R\int 3xy\,dA.$$

Solution From Example 2, you can use the following change of variables.

$$x = \frac{1}{3}(2u + v) \quad \text{and} \quad y = \frac{1}{3}(u - v)$$

The partial derivatives of x and y are

$$\frac{\partial x}{\partial u} = \frac{2}{3}, \quad \frac{\partial x}{\partial v} = \frac{1}{3}, \quad \frac{\partial y}{\partial u} = \frac{1}{3}, \quad \text{and} \quad \frac{\partial y}{\partial v} = -\frac{1}{3}$$

which implies that the Jacobian is

$$\begin{aligned}
\frac{\partial(x, y)}{\partial(u, v)} &= \begin{vmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial x}{\partial v} \\[2mm] \dfrac{\partial y}{\partial u} & \dfrac{\partial y}{\partial v} \end{vmatrix} \\
&= \begin{vmatrix} \dfrac{2}{3} & \dfrac{1}{3} \\[2mm] \dfrac{1}{3} & -\dfrac{1}{3} \end{vmatrix} \\
&= -\frac{2}{9} - \frac{1}{9} \\
&= -\frac{1}{3}.
\end{aligned}$$

So, by Theorem 14.5, you obtain

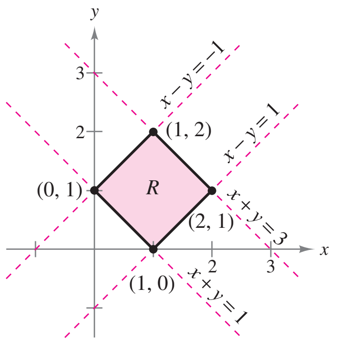
$$\begin{aligned}
\int_R\int 3xy\,dA &= \int_S\int 3\left[\frac{1}{3}(2u + v)\frac{1}{3}(u - v)\right]\left|\frac{\partial(x, y)}{\partial(u, v)}\right| dv\,du \\
&= \int_1^4\int_{-4}^0 \frac{1}{9}(2u^2 - uv - v^2)\,dv\,du \\
&= \frac{1}{9}\int_1^4 \left[2u^2v - \frac{uv^2}{2} - \frac{v^3}{3}\right]_{-4}^0 du \\
&= \frac{1}{9}\int_1^4 \left(8u^2 + 8u - \frac{64}{3}\right) du \\
&= \frac{1}{9}\left[\frac{8u^3}{3} + 4u^2 - \frac{64}{3}u\right]_1^4 \\
&= \frac{164}{9}.
\end{aligned}$$

Try It   Open Exploration

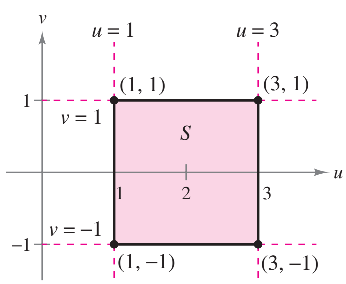
EXAMPLE 4 Using a Change of Variables to Simplify an Integrand

Let R be the region bounded by the square with vertices (0, 1), (1, 2), (2, 1), and (1, 0). Evaluate the integral

$$\int_R\int (x+y)^2 \sin^2(x-y)\, dA.$$



Region R in the xy-plane
Figure 14.76



Region S in the uv-plane
Figure 14.77

Solution Note that the sides of R lie on the lines $x + y = 1$, $x - y = 1$, $x + y = 3$, and $x - y = -1$, as shown in Figure 14.76. Letting $u = x + y$ and $v = x - y$, you can determine the bounds for region S in the uv-plane to be

$$1 \le u \le 3 \qquad \text{and} \qquad -1 \le v \le 1$$

as shown in Figure 14.77. Solving for x and y in terms of u and v produces

$$x = \frac{1}{2}(u+v) \qquad \text{and} \qquad y = \frac{1}{2}(u-v).$$

The partial derivatives of x and y are

$$\frac{\partial x}{\partial u} = \frac{1}{2}, \quad \frac{\partial x}{\partial v} = \frac{1}{2}, \quad \frac{\partial y}{\partial u} = \frac{1}{2}, \quad \text{and} \quad \frac{\partial y}{\partial v} = -\frac{1}{2}$$

which implies that the Jacobian is

$$\frac{\partial(x, y)}{\partial(u, v)} = \begin{vmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial x}{\partial v} \\[2mm] \dfrac{\partial y}{\partial u} & \dfrac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} \dfrac{1}{2} & \dfrac{1}{2} \\[2mm] \dfrac{1}{2} & -\dfrac{1}{2} \end{vmatrix} = -\frac{1}{4} - \frac{1}{4} = -\frac{1}{2}.$$

By Theorem 14.5, it follows that

$$\begin{aligned}
\int_R\int (x+y)^2 \sin^2(x-y)\, dA &= \int_{-1}^{1}\int_{1}^{3} u^2 \sin^2 v\left(\frac{1}{2}\right) du\, dv \\
&= \frac{1}{2}\int_{-1}^{1} (\sin^2 v) \frac{u^3}{3}\Big]_1^3\, dv \\
&= \frac{13}{3}\int_{-1}^{1} \sin^2 v\, dv \\
&= \frac{13}{6}\int_{-1}^{1} (1 - \cos 2v)\, dv \\
&= \frac{13}{6}\left[v - \frac{1}{2}\sin 2v\right]_{-1}^{1} \\
&= \frac{13}{6}\left[2 - \frac{1}{2}\sin 2 + \frac{1}{2}\sin(-2)\right] \\
&= \frac{13}{6}(2 - \sin 2) \\
&\approx 2.363.
\end{aligned}$$

Try It **Exploration A**

In each of the change-of-variables examples in this section, the region S has been a rectangle with sides parallel to the u- or v-axis. Occasionally, a change of variables can be used for other types of regions. For instance, letting $T(u, v) = \left(x, \frac{1}{2}y\right)$ changes the circular region $u^2 + v^2 = 1$ to the elliptical region $x^2 + (y^2/4) = 1$.

Section 15.1 Vector Fields

- **Understand the concept of a vector field.**
- **Determine whether a vector field is conservative.**
- **Find the curl of a vector field.**
- **Find the divergence of a vector field.**

Vector Fields

In Chapter 12, you studied vector-valued functions—functions that assign a vector to a *real number*. There you saw that vector-valued functions of real numbers are useful in representing curves and motion along a curve. In this chapter, you will study two other types of vector-valued functions—functions that assign a vector to a *point in the plane* or a *point in space*. Such functions are called **vector fields,** and they are useful in representing various types of **force fields** and **velocity fields.**

NOTE Although a vector field consists of infinitely many vectors, you can get a good idea of what the vector field looks like by sketching several representative vectors $\mathbf{F}(x, y)$ whose initial points are (x, y).

Definition of Vector Field

Let M and N be functions of two variables x and y, defined on a plane region R. The function $\mathbf{F}$ defined by

$$\mathbf{F}(x, y) = M\mathbf{i} + N\mathbf{j} \qquad \text{Plane}$$

is called a **vector field over *R*.**

Let M, N, and P be functions of three variables x, y, and z, defined on a solid region Q in space. The function $\mathbf{F}$ defined by

$$\mathbf{F}(x, y, z) = M\mathbf{i} + N\mathbf{j} + P\mathbf{k} \qquad \text{Space}$$

is called a **vector field over *Q*.**

From this definition you can see that the *gradient* is one example of a vector field. For example, if

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

then the gradient of f

$$\nabla f(x, y) = f_x(x, y)\mathbf{i} + f_y(x, y)\mathbf{j} = 2x\mathbf{i} + 2y\mathbf{j} \qquad \text{Vector field in the plane}$$

is a vector field in the plane. From Chapter 13, the graphical interpretation of this field is a family of vectors, each of which points in the direction of maximum increase along the surface given by $z = f(x, y)$. For this particular function, the surface is a paraboloid and the gradient tells you that the direction of maximum increase along the surface is the direction given by the ray from the origin through the point (x, y).

Similarly, if

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

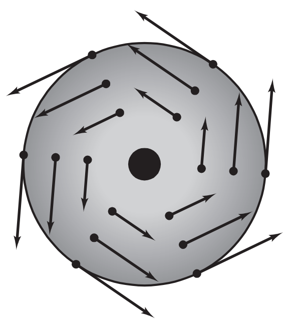
then the gradient of f

$$\begin{aligned}\nabla f(x, y, z) &= f_x(x, y, z)\mathbf{i} + f_y(x, y, z)\mathbf{j} + f_z(x, y, z)\mathbf{k} \\ &= 2x\mathbf{i} + 2y\mathbf{j} + 2z\mathbf{k} \qquad \text{Vector field in space}\end{aligned}$$

is a vector field in space.

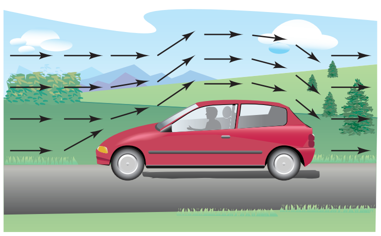
A vector field is **continuous** at a point if each of its component functions M, N, and P is continuous at that point.

Some common *physical* examples of vector fields are **velocity fields, gravitational fields,** and **electric force fields.**



Rotating wheel
Figure 15.1

1. *Velocity fields* describe the motions of systems of particles in the plane or in space. For instance, Figure 15.1 shows the vector field determined by a wheel rotating on an axle. Notice that the velocity vectors are determined by the locations of their initial points—the farther a point is from the axle, the greater its velocity. Velocity fields are also determined by the flow of liquids through a container or by the flow of air currents around a moving object, as shown in Figure 15.2.


Air flow vector field
Figure 15.2

2. *Gravitational fields* are defined by **Newton's Law of Gravitation,** which states that the force of attraction exerted on a particle of mass m_1 located at (x, y, z) by a particle of mass m_2 located at $(0, 0, 0)$ is given by

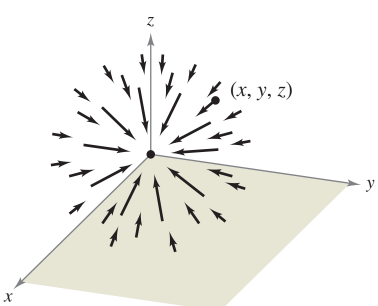
$$\mathbf{F}(x, y, z) = \frac{-Gm_1m_2}{x^2 + y^2 + z^2}\mathbf{u}$$

where G is the gravitational constant and $\mathbf{u}$ is the unit vector in the direction from the origin to (x, y, z). In Figure 15.3, you can see that the gravitational field $\mathbf{F}$ has the properties that $\mathbf{F}(x, y, z)$ always points toward the origin, and that the magnitude of $\mathbf{F}(x, y, z)$ is the same at all points equidistant from the origin. A vector field with these two properties is called a **central force field.** Using the position vector

$$\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$$

for the point (x, y, z), you can write the gravitational field $\mathbf{F}$ as

$$\begin{aligned}\mathbf{F}(x, y, z) &= \frac{-Gm_1m_2}{\|\mathbf{r}\|^2}\left(\frac{\mathbf{r}}{\|\mathbf{r}\|}\right)\\ &= \frac{-Gm_1m_2}{\|\mathbf{r}\|^2}\mathbf{u}.\end{aligned}$$



Gravitational force field
Figure 15.3

3. *Electric force fields* are defined by **Coulomb's Law,** which states that the force exerted on a particle with electric charge q_1 located at (x, y, z) by a particle with electric charge q_2 located at $(0, 0, 0)$ is given by

$$\mathbf{F}(x, y, z) = \frac{cq_1q_2}{\|\mathbf{r}\|^2}\mathbf{u}$$

where $\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$, $\mathbf{u} = \mathbf{r}/\|\mathbf{r}\|$, and c is a constant that depends on the choice of units for $\|\mathbf{r}\|$, q_1, and q_2.

Note that an electric force field has the same form as a gravitational field. That is,

$$\mathbf{F}(x, y, z) = \frac{k}{\|\mathbf{r}\|^2}\mathbf{u}.$$

Such a force field is called an **inverse square field.**

Definition of Inverse Square Field

Let $\mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j} + z(t)\mathbf{k}$ be a position vector. The vector field $\mathbf{F}$ is an **inverse square field** if

$$\mathbf{F}(x, y, z) = \frac{k}{\|\mathbf{r}\|^2}\mathbf{u}$$

where k is a real number and $\mathbf{u} = \mathbf{r}/\|\mathbf{r}\|$ is a unit vector in the direction of $\mathbf{r}$.

Because vector fields consist of infinitely many vectors, it is not possible to create a sketch of the entire field. Instead, when you sketch a vector field, your goal is to sketch representative vectors that help you visualize the field.

EXAMPLE 1 Sketching a Vector Field

Sketch some vectors in the vector field given by

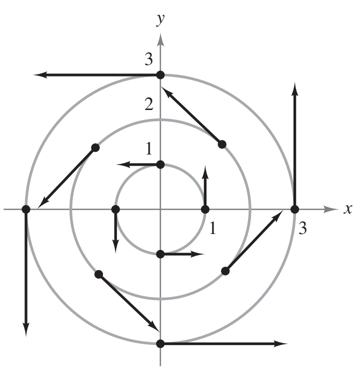
$$\mathbf{F}(x, y) = -y\mathbf{i} + x\mathbf{j}.$$

Solution You could plot vectors at several random points in the plane. However, it is more enlightening to plot vectors of equal magnitude. This corresponds to finding level curves in scalar fields. In this case, vectors of equal magnitude lie on circles.

$$\|\mathbf{F}\| = c \qquad \text{Vectors of length } c$$
$$\sqrt{x^2 + y^2} = c$$
$$x^2 + y^2 = c^2 \qquad \text{Equation of circle}$$

To begin making the sketch, choose a value for c and plot several vectors on the resulting circle. For instance, the following vectors occur on the unit circle.

Point	*Vector*
$(1, 0)$	$\mathbf{F}(1, 0) = \mathbf{j}$
$(0, 1)$	$\mathbf{F}(0, 1) = -\mathbf{i}$
$(-1, 0)$	$\mathbf{F}(-1, 0) = -\mathbf{j}$
$(0, -1)$	$\mathbf{F}(0, -1) = \mathbf{i}$



Vector field:
$\mathbf{F}(x, y) = -y\mathbf{i} + x\mathbf{j}$

Figure 15.4

These and several other vectors in the vector field are shown in Figure 15.4. Note in the figure that this vector field is similar to that given by the rotating wheel shown in Figure 15.1.

Try It Exploration A

EXAMPLE 2 Sketching a Vector Field

Sketch some vectors in the vector field given by

$$\mathbf{F}(x, y) = 2x\mathbf{i} + y\mathbf{j}.$$

Solution For this vector field, vectors of equal length lie on ellipses given by

$$\|\mathbf{F}\| = \sqrt{(2x)^2 + (y)^2} = c$$

which implies that

$$4x^2 + y^2 = c^2.$$

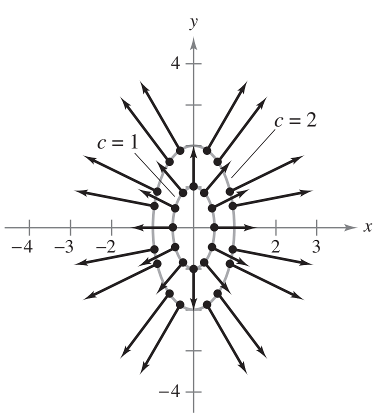
For $c = 1$, sketch several vectors $2x\mathbf{i} + y\mathbf{j}$ of magnitude 1 at points on the ellipse given by

$$4x^2 + y^2 = 1.$$

For $c = 2$, sketch several vectors $2x\mathbf{i} + y\mathbf{j}$ of magnitude 2 at points on the ellipse given by

$$4x^2 + y^2 = 4.$$

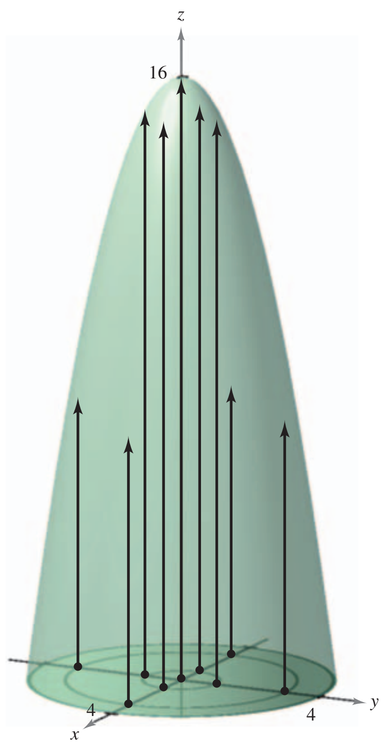
These vectors are shown in Figure 15.5.



Vector field:
$\mathbf{F}(x, y) = 2x\mathbf{i} + y\mathbf{j}$

Figure 15.5

Try It Exploration A



Velocity field:
$\mathbf{v}(x, y, z) = (16 - x^2 - y^2)\mathbf{k}$

Figure 15.6

Rotatable Graph

EXAMPLE 3 Sketching a Velocity Field

Sketch some vectors in the velocity field given by

$$\mathbf{v}(x, y, z) = (16 - x^2 - y^2)\mathbf{k}$$

where $x^2 + y^2 \leq 16$.

Solution You can imagine that $\mathbf{v}$ describes the velocity of a liquid flowing through a tube of radius 4. Vectors near the z-axis are longer than those near the edge of the tube. For instance, at the point $(0, 0, 0)$, the velocity vector is $\mathbf{v}(0, 0, 0) = 16\mathbf{k}$, whereas at the point $(0, 3, 0)$, the velocity vector is $\mathbf{v}(0, 3, 0) = 7\mathbf{k}$. Figure 15.6 shows these and several other vectors for the velocity field. From the figure, you can see that the speed of the liquid is greater near the center of the tube than near the edges of the tube.

Try It   Exploration A

Conservative Vector Fields

Notice in Figure 15.5 that all the vectors appear to be normal to the level curve from which they emanate. Because this is a property of gradients, it is natural to ask whether the vector field given by $\mathbf{F}(x, y) = 2x\mathbf{i} + y\mathbf{j}$ is the *gradient* for some differentiable function f. The answer is that some vector fields can be represented as the gradients of differentiable functions and some cannot—those that can are called **conservative** vector fields.

Definition of Conservative Vector Field

A vector field $\mathbf{F}$ is called **conservative** if there exists a differentiable function f such that $\mathbf{F} = \nabla f$. The function f is called the **potential function** for $\mathbf{F}$.

EXAMPLE 4 Conservative Vector Fields

a. The vector field given by $\mathbf{F}(x, y) = 2x\mathbf{i} + y\mathbf{j}$ is conservative. To see this, consider the potential function $f(x, y) = x^2 + \frac{1}{2}y^2$. Because

$$\nabla f = 2x\mathbf{i} + y\mathbf{j} = \mathbf{F}$$

it follows that $\mathbf{F}$ is conservative.

b. Every inverse square field is conservative. To see this, let

$$\mathbf{F}(x, y, z) = \frac{k}{\|\mathbf{r}\|^2}\mathbf{u} \quad \text{and} \quad f(x, y, z) = \frac{-k}{\sqrt{x^2 + y^2 + z^2}}$$

where $\mathbf{u} = \mathbf{r}/\|\mathbf{r}\|$. Because

$$\begin{aligned}\nabla f &= \frac{kx}{(x^2 + y^2 + z^2)^{3/2}}\mathbf{i} + \frac{ky}{(x^2 + y^2 + z^2)^{3/2}}\mathbf{j} + \frac{kz}{(x^2 + y^2 + z^2)^{3/2}}\mathbf{k} \\ &= \frac{k}{x^2 + y^2 + z^2}\left(\frac{x\mathbf{i} + y\mathbf{j} + z\mathbf{k}}{\sqrt{x^2 + y^2 + z^2}}\right) \\ &= \frac{k}{\|\mathbf{r}\|^2}\frac{\mathbf{r}}{\|\mathbf{r}\|} \\ &= \frac{k}{\|\mathbf{r}\|^2}\mathbf{u}\end{aligned}$$

it follows that $\mathbf{F}$ is conservative.

Try It   Exploration A

As can be seen in Example 4(b), many important vector fields, including gravitational fields and electric force fields, are conservative. Most of the terminology in this chapter comes from physics. For example, the term "conservative" is derived from the classic physical law regarding the conservation of energy. This law states that the sum of the kinetic energy and the potential energy of a particle moving in a conservative force field is constant. (The kinetic energy of a particle is the energy due to its motion, and the potential energy is the energy due to its position in the force field.)

The following important theorem gives a necessary and sufficient condition for a vector field *in the plane* to be conservative.

THEOREM 15.1 Test for Conservative Vector Field in the Plane

Let M and N have continuous first partial derivatives on an open disk R. The vector field given by $\mathbf{F}(x, y) = M\mathbf{i} + N\mathbf{j}$ is conservative if and only if

$$\frac{\partial N}{\partial x} = \frac{\partial M}{\partial y}.$$

Proof To prove that the given condition is necessary for $\mathbf{F}$ to be conservative, suppose there exists a potential function f such that

$$\mathbf{F}(x, y) = \nabla f(x, y) = M\mathbf{i} + N\mathbf{j}.$$

Then you have

$$f_x(x, y) = M \quad \Longrightarrow \quad f_{xy}(x, y) = \frac{\partial M}{\partial y}$$

$$f_y(x, y) = N \quad \Longrightarrow \quad f_{yx}(x, y) = \frac{\partial N}{\partial x}$$

and, by the equivalence of the mixed partials f_{xy} and f_{yx}, you can conclude that $\partial N/\partial x = \partial M/\partial y$ for all (x, y) in R. The sufficiency of the condition is proved in Section 15.4.

NOTE Theorem 15.1 requires that the domain of $\mathbf{F}$ be an open disk. If R is simply an open region, the given condition is necessary but not sufficient to produce a conservative vector field.

EXAMPLE 5 Testing for Conservative Vector Fields in the Plane

Decide whether the vector field given by $\mathbf{F}$ is conservative.

a. $\mathbf{F}(x, y) = x^2y\mathbf{i} + xy\mathbf{j}$ **b.** $\mathbf{F}(x, y) = 2x\mathbf{i} + y\mathbf{j}$

Solution

a. The vector field given by $\mathbf{F}(x, y) = x^2y\mathbf{i} + xy\mathbf{j}$ is not conservative because

$$\frac{\partial M}{\partial y} = \frac{\partial}{\partial y}[x^2y] = x^2 \quad \text{and} \quad \frac{\partial N}{\partial x} = \frac{\partial}{\partial x}[xy] = y.$$

b. The vector field given by $\mathbf{F}(x, y) = 2x\mathbf{i} + y\mathbf{j}$ is conservative because

$$\frac{\partial M}{\partial y} = \frac{\partial}{\partial y}[2x] = 0 \quad \text{and} \quad \frac{\partial N}{\partial x} = \frac{\partial}{\partial x}[y] = 0.$$

Try It Exploration A

Theorem 15.1 tells you whether a vector field is conservative. It does not tell you how to find a potential function of **F**. The problem is comparable to antidifferentiation. Sometimes you will be able to find a potential function by simple inspection. For instance, in Example 4 you observed that

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

has the property that $\nabla f(x, y) = 2x\mathbf{i} + y\mathbf{j}$.

EXAMPLE 6 Finding a Potential Function for F(x, y)

Find a potential function for

$$\mathbf{F}(x, y) = 2xy\mathbf{i} + (x^2 - y)\mathbf{j}.$$

Solution From Theorem 15.1 it follows that **F** is conservative because

$$\frac{\partial}{\partial y}[2xy] = 2x \quad \text{and} \quad \frac{\partial}{\partial x}[x^2 - y] = 2x.$$

If f is a function whose gradient is equal to $\mathbf{F}(x, y)$, then

$$\nabla f(x, y) = 2xy\mathbf{i} + (x^2 - y)\mathbf{j}$$

which implies that

$$f_x(x, y) = 2xy$$

and

$$f_y(x, y) = x^2 - y.$$

To reconstruct the function f from these two partial derivatives, integrate $f_x(x, y)$ with respect to x and $f_y(x, y)$ with respect to y, as follows.

$$f(x, y) = \int f_x(x, y)\, dx = \int 2xy\, dx = x^2y + g(y)$$

$$f(x, y) = \int f_y(x, y)\, dy = \int (x^2 - y)\, dy = x^2y - \frac{y^2}{2} + h(x)$$

Notice that $g(y)$ is constant with respect to x and $h(x)$ is constant with respect to y. To find a single expression that represents $f(x, y)$, let

$$g(y) = -\frac{y^2}{2} \quad \text{and} \quad h(x) = K.$$

Then, you can write

$$\begin{aligned} f(x, y) &= x^2y + g(y) + K \\ &= x^2y - \frac{y^2}{2} + K. \end{aligned}$$

You can check this result by forming the gradient of f. You will see that it is equal to the original function **F**.

Try It | Exploration A | Exploration B

NOTE Notice that the solution in Example 6 is comparable to that given by an indefinite integral. That is, the solution represents a family of potential functions, any two of which differ by a constant. To find a unique solution, you would have to be given an initial condition satisfied by the potential function.

Curl of a Vector Field

Theorem 15.1 has a counterpart for vector fields in space. Before stating that result, the definition of the **curl of a vector field** in space is given.

Definition of Curl of a Vector Field

The curl of $\mathbf{F}(x, y, z) = M\mathbf{i} + N\mathbf{j} + P\mathbf{k}$ is

$$\begin{aligned}\textbf{curl } \mathbf{F}(x, y, z) &= \nabla \times \mathbf{F}(x, y, z)\\ &= \left(\frac{\partial P}{\partial y} - \frac{\partial N}{\partial z}\right)\mathbf{i} - \left(\frac{\partial P}{\partial x} - \frac{\partial M}{\partial z}\right)\mathbf{j} + \left(\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y}\right)\mathbf{k}.\end{aligned}$$

NOTE If **curl F** $= \mathbf{0}$, then **F** is said to be **irrotational.**

The cross product notation used for curl comes from viewing the gradient ∇f as the result of the **differential operator** ∇ acting on the function f. In this context, you can use the following determinant form as an aid in remembering the formula for curl.

$$\begin{aligned}\textbf{curl } \mathbf{F}(x, y, z) &= \nabla \times \mathbf{F}(x, y, z)\\ &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ M & N & P \end{vmatrix}\\ &= \left(\frac{\partial P}{\partial y} - \frac{\partial N}{\partial z}\right)\mathbf{i} - \left(\frac{\partial P}{\partial x} - \frac{\partial M}{\partial z}\right)\mathbf{j} + \left(\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y}\right)\mathbf{k}\end{aligned}$$

EXAMPLE 7 Finding the Curl of a Vector Field

Find **curl F** for the vector field given by

$$\mathbf{F}(x, y, z) = 2xy\mathbf{i} + (x^2 + z^2)\mathbf{j} + 2yz\mathbf{k}.$$

Is **F** irrotational?

Solution The curl of **F** is given by

$$\begin{aligned}\textbf{curl } \mathbf{F}(x, y, z) &= \nabla \times \mathbf{F}(x, y, z)\\ &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ 2xy & x^2 + z^2 & 2yz \end{vmatrix}\\ &= \begin{vmatrix} \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ x^2 + z^2 & 2yz \end{vmatrix}\mathbf{i} - \begin{vmatrix} \frac{\partial}{\partial x} & \frac{\partial}{\partial z} \\ 2xy & 2yz \end{vmatrix}\mathbf{j} + \begin{vmatrix} \frac{\partial}{\partial x} & \frac{\partial}{\partial y} \\ 2xy & x^2 + z^2 \end{vmatrix}\mathbf{k}\\ &= (2z - 2z)\mathbf{i} - (0 - 0)\mathbf{j} + (2x - 2x)\mathbf{k}\\ &= \mathbf{0}.\end{aligned}$$

Because **curl F** $= \mathbf{0}$, **F** is irrotational.

Try It Open Exploration

Later in this chapter, you will assign a physical interpretation to the curl of a vector field. But for now, the primary use of curl is shown in the following test for conservative vector fields in space. The test states that for a vector field whose domain is all of three-dimensional space (or an open sphere), the curl is **0** at every point in the domain if and only if **F** is conservative. The proof is similar to that given for Theorem 15.1.

THEOREM 15.2 Test for Conservative Vector Field in Space

Suppose that M, N, and P have continuous first partial derivatives in an open sphere Q in space. The vector field given by $\mathbf{F}(x, y, z) = M\mathbf{i} + N\mathbf{j} + P\mathbf{k}$ is conservative if and only if

$$\mathbf{curl}\ \mathbf{F}(x, y, z) = \mathbf{0}.$$

That is, **F** is conservative if and only if

$$\frac{\partial P}{\partial y} = \frac{\partial N}{\partial z}, \quad \frac{\partial P}{\partial x} = \frac{\partial M}{\partial z}, \quad \text{and} \quad \frac{\partial N}{\partial x} = \frac{\partial M}{\partial y}.$$

From Theorem 15.2, you can see that the vector field given in Example 7 is conservative because $\mathbf{curl}\ \mathbf{F}(x, y, z) = \mathbf{0}$. Try showing that the vector field

$$\mathbf{F}(x, y, z) = x^3y^2z\mathbf{i} + x^2z\mathbf{j} + x^2y\mathbf{k}$$

is not conservative—you can do this by showing that its curl is

$$\mathbf{curl}\ \mathbf{F}(x, y, z) = (x^3y^2 - 2xy)\mathbf{j} + (2xz - 2x^3yz)\mathbf{k} \neq \mathbf{0}.$$

For vector fields in space that pass the test for being conservative, you can find a potential function by following the same pattern used in the plane (as demonstrated in Example 6).

EXAMPLE 8 Finding a Potential Function for F(*x*, *y*, *z*)

NOTE Examples 6 and 8 are illustrations of a type of problem called *recovering a function from its gradient*. If you go on to take a course in differential equations, you will study other methods for solving this type of problem. One popular method gives an interplay between successive "partial integrations" and partial differentiations.

Find a potential function for $\mathbf{F}(x, y, z) = 2xy\mathbf{i} + (x^2 + z^2)\mathbf{j} + 2yz\mathbf{k}$.

Solution From Example 7, you know that the vector field given by **F** is conservative. If f is a function such that $\mathbf{F}(x, y, z) = \nabla f(x, y, z)$, then

$$f_x(x, y, z) = 2xy, \quad f_y(x, y, z) = x^2 + z^2, \quad \text{and} \quad f_z(x, y, z) = 2yz$$

and integrating with respect to x, y, and z separately produces

$$f(x, y, z) = \int M\,dx = \int 2xy\,dx = x^2y + g(y, z)$$

$$f(x, y, z) = \int N\,dy = \int (x^2 + z^2)\,dy = x^2y + yz^2 + h(x, z)$$

$$f(x, y, z) = \int P\,dz = \int 2yz\,dz = yz^2 + k(x, y).$$

Comparing these three versions of $f(x, y, z)$, you can conclude that

$$g(y, z) = yz^2 + K, \quad h(x, z) = K, \quad \text{and} \quad k(x, y) = x^2y + K.$$

So, $f(x, y, z)$ is given by

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

Try It Exploration A

Divergence of a Vector Field

NOTE Divergence can be viewed as a type of derivative of **F** in that, for vector fields representing velocities of moving particles, the divergence measures the rate of particle flow per unit volume at a point. In hydrodynamics (the study of fluid motion), a velocity field that is divergence free is called **incompressible.** In the study of electricity and magnetism, a vector field that is divergence free is called **solenoidal.**

You have seen that the curl of a vector field **F** is itself a vector field. Another important function defined on a vector field is **divergence,** which is a scalar function.

Definition of Divergence of a Vector Field

The **divergence** of $\mathbf{F}(x, y) = M\mathbf{i} + N\mathbf{j}$ is

$$\text{div } \mathbf{F}(x, y) = \nabla \cdot \mathbf{F}(x, y) = \frac{\partial M}{\partial x} + \frac{\partial N}{\partial y}. \qquad \text{Plane}$$

The **divergence** of $\mathbf{F}(x, y, z) = M\mathbf{i} + N\mathbf{j} + P\mathbf{k}$ is

$$\text{div } \mathbf{F}(x, y, z) = \nabla \cdot \mathbf{F}(x, y, z) = \frac{\partial M}{\partial x} + \frac{\partial N}{\partial y} + \frac{\partial P}{\partial z}. \qquad \text{Space}$$

If $\text{div } \mathbf{F} = 0$, then **F** is said to be **divergence free.**

The dot product notation used for divergence comes from considering ∇ as a **differential operator,** as follows.

$$\begin{aligned}\nabla \cdot \mathbf{F}(x, y, z) &= \left[\left(\frac{\partial}{\partial x}\right)\mathbf{i} + \left(\frac{\partial}{\partial y}\right)\mathbf{j} + \left(\frac{\partial}{\partial z}\right)\mathbf{k}\right] \cdot (M\mathbf{i} + N\mathbf{j} + P\mathbf{k}) \\ &= \frac{\partial M}{\partial x} + \frac{\partial N}{\partial y} + \frac{\partial P}{\partial z}\end{aligned}$$

EXAMPLE 9 Finding the Divergence of a Vector Field

Find the divergence at $(2, 1, -1)$ for the vector field

$$\mathbf{F}(x, y, z) = x^3y^2z\mathbf{i} + x^2z\mathbf{j} + x^2y\mathbf{k}.$$

Solution The divergence of **F** is

$$\text{div } \mathbf{F}(x, y, z) = \frac{\partial}{\partial x}[x^3y^2z] + \frac{\partial}{\partial y}[x^2z] + \frac{\partial}{\partial z}[x^2y] = 3x^2y^2z.$$

At the point $(2, 1, -1)$, the divergence is

$$\text{div } \mathbf{F}(2, 1, -1) = 3(2^2)(1^2)(-1) = -12.$$

Try It

Exploration A

There are many important properties of the divergence and curl of a vector field **F** (see Exercises 77–83). One that is used often is described in Theorem 15.3. You are asked to prove this theorem in Exercise 84.

THEOREM 15.3 Relationship Between Divergence and Curl

If $\mathbf{F}(x, y, z) = M\mathbf{i} + N\mathbf{j} + P\mathbf{k}$ is a vector field and M, N, and P have continuous second partial derivatives, then

$$\text{div } (\mathbf{curl}\ \mathbf{F}) = 0.$$

Section 15.2 Line Integrals

- **Understand and use the concept of a piecewise smooth curve.**
- **Write and evaluate a line integral.**
- **Write and evaluate a line integral of a vector field.**
- **Write and evaluate a line integral in differential form.**

Piecewise Smooth Curves

A classic property of gravitational fields is that, subject to certain physical constraints, the work done by gravity on an object moving between two points in the field is independent of the path taken by the object. One of the constraints is that the **path** must be a piecewise smooth curve. Recall that a plane curve C given by

$$\mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j}, \quad a \le t \le b$$

is **smooth** if

$$\frac{dx}{dt} \quad \text{and} \quad \frac{dy}{dt}$$

are continuous on $[a, b]$ and not simultaneously 0 on (a, b). Similarly, a space curve C given by

$$\mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j} + z(t)\mathbf{k}, \quad a \le t \le b$$

is **smooth** if

$$\frac{dx}{dt}, \quad \frac{dy}{dt}, \quad \text{and} \quad \frac{dz}{dt}$$

are continuous on $[a, b]$ and not simultaneously 0 on (a, b). A curve C is **piecewise smooth** if the interval $[a, b]$ can be partitioned into a finite number of subintervals, on each of which C is smooth.

> **JOSIAH WILLARD GIBBS (1839–1903)**
>
> Many physicists and mathematicians have contributed to the theory and applications described in this chapter—Newton, Gauss, Laplace, Hamilton, and Maxwell, among others. However, the use of vector analysis to describe these results is attributed primarily to the American mathematical physicist Josiah Willard Gibbs.

EXAMPLE 1 Finding a Piecewise Smooth Parametrization

Find a piecewise smooth parametrization of the graph of C shown in Figure 15.7.

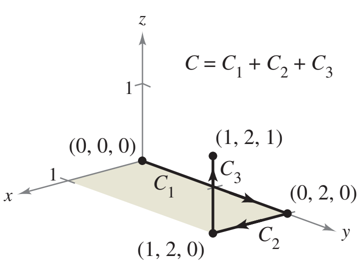


Figure 15.7

Solution Because C consists of three line segments C_1, C_2, and C_3, you can construct a smooth parametrization for each segment and piece them together by making the last t-value in C_i correspond to the first t-value in C_{i+1}, as follows.

$$\begin{array}{llll} C_1\colon\ x(t) = 0, & y(t) = 2t, & z(t) = 0, & 0 \le t \le 1 \\ C_2\colon\ x(t) = t - 1, & y(t) = 2, & z(t) = 0, & 1 \le t \le 2 \\ C_3\colon\ x(t) = 1, & y(t) = 2, & z(t) = t - 2, & 2 \le t \le 3 \end{array}$$

So, C is given by

$$\mathbf{r}(t) = \begin{cases} 2t\mathbf{j}, & 0 \le t \le 1 \\ (t-1)\mathbf{i} + 2\mathbf{j}, & 1 \le t \le 2 \\ \mathbf{i} + 2\mathbf{j} + (t-2)\mathbf{k}, & 2 \le t \le 3 \end{cases}.$$

Because C_1, C_2, and C_3 are smooth, it follows that C is piecewise smooth.

Try It **Exploration A**

Recall that parametrization of a curve induces an **orientation** to the curve. For instance, in Example 1, the curve is oriented such that the positive direction is from $(0, 0, 0)$, following the curve to $(1, 2, 1)$. Try finding a parametrization that induces the opposite orientation.

Line Integrals

Up to this point in the text, you have studied various types of integrals. For a single integral

$$\int_a^b f(x)\,dx \qquad \text{Integrate over interval } [a, b].$$

you integrated over the interval $[a, b]$. Similarly, for a double integral

$$\int\!\!\int_R f(x, y)\,dA \qquad \text{Integrate over region } R.$$

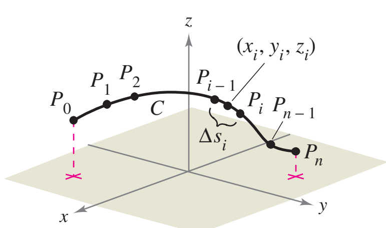
you integrated over the region R in the plane. In this section, you will study a new type of integral called a **line integral**

$$\int_C f(x, y)\,ds \qquad \text{Integrate over curve } C.$$

for which you integrate over a piecewise smooth curve C. (The terminology is somewhat unfortunate—this type of integral might be better described as a "curve integral.")

To introduce the concept of a line integral, consider the mass of a wire of finite length, given by a curve C in space. The density (mass per unit length) of the wire at the point (x, y, z) is given by $f(x, y, z)$. Partition the curve C by the points

$$P_0, P_1, \ldots, P_n$$



Partitioning of curve C
Figure 15.8

producing n subarcs, as shown in Figure 15.8. The length of the ith subarc is given by Δs_i. Next, choose a point (x_i, y_i, z_i) in each subarc. If the length of each subarc is small, the total mass of the wire can be approximated by the sum

$$\text{Mass of wire} \approx \sum_{i=1}^{n} f(x_i, y_i, z_i)\,\Delta s_i.$$

If you let $\|\Delta\|$ denote the length of the longest subarc and let $\|\Delta\|$ approach 0, it seems reasonable that the limit of this sum approaches the mass of the wire. This leads to the following definition.

Definition of Line Integral

If f is defined in a region containing a smooth curve C of finite length, then the **line integral of f along C** is given by

$$\int_C f(x, y)\,ds = \lim_{\|\Delta\|\to 0} \sum_{i=1}^{n} f(x_i, y_i)\,\Delta s_i \qquad \text{Plane}$$

or

$$\int_C f(x, y, z)\,ds = \lim_{\|\Delta\|\to 0} \sum_{i=1}^{n} f(x_i, y_i, z_i)\,\Delta s_i \qquad \text{Space}$$

provided this limit exists.

As with the integrals discussed in Chapter 14, evaluation of a line integral is best accomplished by converting to a definite integral. It can be shown that if f is *continuous*, the limit given above exists and is the same for all smooth parametrizations of C.

To evaluate a line integral over a plane curve C given by $\mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j}$, use the fact that

$$ds = \|\mathbf{r}'(t)\|\,dt = \sqrt{[x'(t)]^2 + [y'(t)]^2}\,dt.$$

A similar formula holds for a space curve, as indicated in Theorem 15.4.

THEOREM 15.4 Evaluation of a Line Integral as a Definite Integral

Let f be continuous in a region containing a smooth curve C. If C is given by $\mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j}$, where $a \le t \le b$, then

$$\int_C f(x, y)\,ds = \int_a^b f(x(t), y(t))\sqrt{[x'(t)]^2 + [y'(t)]^2}\,dt.$$

If C is given by $\mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j} + z(t)\mathbf{k}$, where $a \le t \le b$, then

$$\int_C f(x, y, z)\,ds = \int_a^b f(x(t), y(t), z(t))\sqrt{[x'(t)]^2 + [y'(t)]^2 + [z'(t)]^2}\,dt.$$

Note that if $f(x, y, z) = 1$, the line integral gives the arc length of the curve C, as defined in Section 12.5. That is,

$$\int_C 1\,ds = \int_a^b \|\mathbf{r}'(t)\|\,dt = \text{length of curve } C.$$

EXAMPLE 2 Evaluating a Line Integral

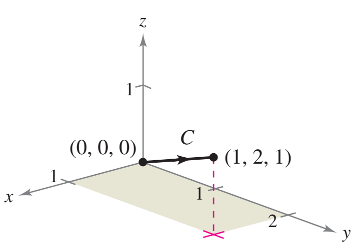


Figure 15.9

Evaluate

$$\int_C (x^2 - y + 3z)\,ds$$

where C is the line segment shown in Figure 15.9.

Solution Begin by writing a parametric form of the equation of a line:

$$x = t, \quad y = 2t, \quad \text{and} \quad z = t, \quad 0 \le t \le 1.$$

Therefore, $x'(t) = 1$, $y'(t) = 2$, and $z'(t) = 1$, which implies that

$$\sqrt{[x'(t)]^2 + [y'(t)]^2 + [z'(t)]^2} = \sqrt{1^2 + 2^2 + 1^2} = \sqrt{6}.$$

So, the line integral takes the following form.

$$\begin{aligned}\int_C (x^2 - y + 3z)\,ds &= \int_0^1 (t^2 - 2t + 3t)\sqrt{6}\,dt \\ &= \sqrt{6}\int_0^1 (t^2 + t)\,dt \\ &= \sqrt{6}\left[\frac{t^3}{3} + \frac{t^2}{2}\right]_0^1 \\ &= \frac{5\sqrt{6}}{6}\end{aligned}$$

NOTE The value of the line integral in Example 2 does not depend on the parametrization of the line segment C (any smooth parametrization will produce the same value). To convince yourself of this, try some other parametrizations, such as $x = 1 + 2t$, $y = 2 + 4t$, $z = 1 + 2t$, $-\frac{1}{2} \le t \le 0$, or $x = -t$, $y = -2t$, $z = -t$, $-1 \le t \le 0$.

Try It

Exploration A

Suppose C is a path composed of smooth curves $C_1, C_2, \ldots, C_n$. If f is continuous on C, it can be shown that

$$\int_C f(x, y)\, ds = \int_{C_1} f(x, y)\, ds + \int_{C_2} f(x, y)\, ds + \cdots + \int_{C_n} f(x, y)\, ds.$$

This property is used in Example 3.

EXAMPLE 3 Evaluating a Line Integral Over a Path

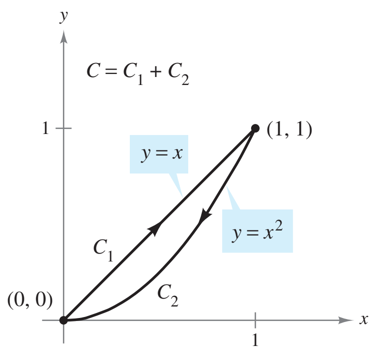


Figure 15.10

Evaluate $\int_C x\, ds$, where C is the piecewise smooth curve shown in Figure 15.10.

Solution Begin by integrating up the line $y = x$, using the following parametrization.

$$C_1\colon\ x = t,\ y = t, \quad 0 \le t \le 1$$

For this curve, $\mathbf{r}(t) = t\mathbf{i} + t\mathbf{j}$, which implies that $x'(t) = 1$ and $y'(t) = 1$. So,

$$\sqrt{[x'(t)]^2 + [y'(t)]^2} = \sqrt{2}$$

and you have

$$\int_{C_1} x\, ds = \int_0^1 t\sqrt{2}\, dt = \frac{\sqrt{2}}{2} t^2 \Big]_0^1 = \frac{\sqrt{2}}{2}.$$

Next, integrate down the parabola $y = x^2$, using the parametrization

$$C_2\colon\ x = 1 - t, \quad y = (1 - t)^2, \quad 0 \le t \le 1.$$

For this curve, $\mathbf{r}(t) = (1 - t)\mathbf{i} + (1 - t)^2\mathbf{j}$, which implies that $x'(t) = -1$ and $y'(t) = -2(1 - t)$. So,

$$\sqrt{[x'(t)]^2 + [y'(t)]^2} = \sqrt{1 + 4(1 - t)^2}$$

and you have

$$\begin{aligned}\int_{C_2} x\, ds &= \int_0^1 (1 - t)\sqrt{1 + 4(1 - t)^2}\, dt \\ &= -\frac{1}{8}\left[\frac{2}{3}[1 + 4(1 - t)^2]^{3/2}\right]_0^1 \\ &= \frac{1}{12}(5^{3/2} - 1).\end{aligned}$$

Consequently,

$$\int_C x\, ds = \int_{C_1} x\, ds + \int_{C_2} x\, ds = \frac{\sqrt{2}}{2} + \frac{1}{12}(5^{3/2} - 1) \approx 1.56.$$

Try It Exploration A

For parametrizations given by $\mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j} + z(t)\mathbf{k}$, it is helpful to remember the form of ds as

$$ds = \|\mathbf{r}'(t)\|\, dt = \sqrt{[x'(t)]^2 + [y'(t)]^2 + [z'(t)]^2}\, dt.$$

This is demonstrated in Example 4.

EXAMPLE 4 Evaluating a Line Integral

Evaluate $\displaystyle\int_C (x + 2)\, ds$, where C is the curve represented by

$$\mathbf{r}(t) = t\mathbf{i} + \frac{4}{3}t^{3/2}\mathbf{j} + \frac{1}{2}t^2\mathbf{k}, \quad 0 \le t \le 2.$$

Solution Because $\mathbf{r}'(t) = \mathbf{i} + 2t^{1/2}\mathbf{j} + t\mathbf{k}$, and

$$\|\mathbf{r}'(t)\| = \sqrt{[x'(t)]^2 + [y'(t)]^2 + [z'(t)]^2} = \sqrt{1 + 4t + t^2}$$

it follows that

$$\begin{aligned}
\int_C (x + 2)\, ds &= \int_0^2 (t + 2)\sqrt{1 + 4t + t^2}\, dt \\
&= \frac{1}{2}\int_0^2 2(t + 2)(1 + 4t + t^2)^{1/2}\, dt \\
&= \frac{1}{3}\Big[(1 + 4t + t^2)^{3/2}\Big]_0^2 \\
&= \frac{1}{3}\left(13\sqrt{13} - 1\right) \\
&\approx 15.29.
\end{aligned}$$

Try It | Exploration A

The next example shows how a line integral can be used to find the mass of a spring whose density varies. In Figure 15.11, note that the density of this spring increases as the spring spirals up the z-axis.

EXAMPLE 5 Finding the Mass of a Spring

Find the mass of a spring in the shape of the circular helix

$$\mathbf{r}(t) = \frac{1}{\sqrt{2}}(\cos t\mathbf{i} + \sin t\mathbf{j} + t\mathbf{k}), \quad 0 \le t \le 6\pi$$

where the density of the spring is $\rho(x, y, z) = 1 + z$, as shown in Figure 15.11.

Solution Because

$$\|\mathbf{r}'(t)\| = \frac{1}{\sqrt{2}}\sqrt{(-\sin t)^2 + (\cos t)^2 + (1)^2} = 1$$

it follows that the mass of the spring is

$$\begin{aligned}
\text{Mass} = \int_C (1 + z)\, ds &= \int_0^{6\pi}\left(1 + \frac{t}{\sqrt{2}}\right) dt \\
&= \left[t + \frac{t^2}{2\sqrt{2}}\right]_0^{6\pi} \\
&= 6\pi\left(1 + \frac{3\pi}{\sqrt{2}}\right) \\
&\approx 144.47.
\end{aligned}$$

The mass of the spring is approximately 144.47.

Try It | Exploration A

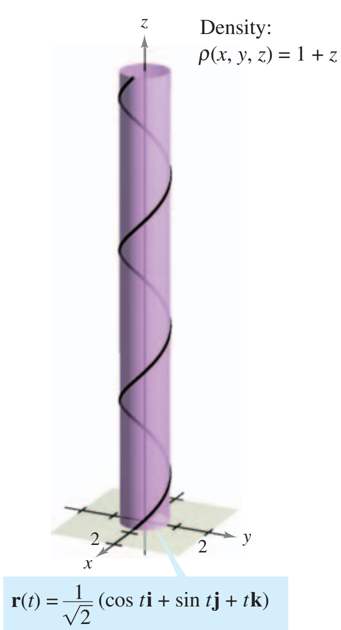
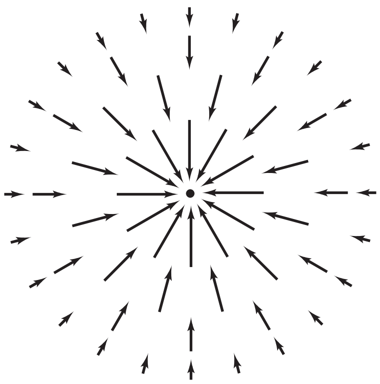
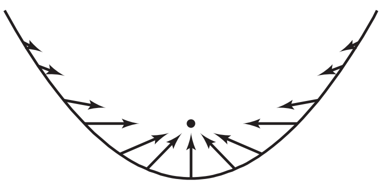


Figure 15.11

Rotatable Graph



Inverse square force field **F**



Vectors along a parabolic path in the force field **F**

Figure 15.12

Line Integrals of Vector Fields

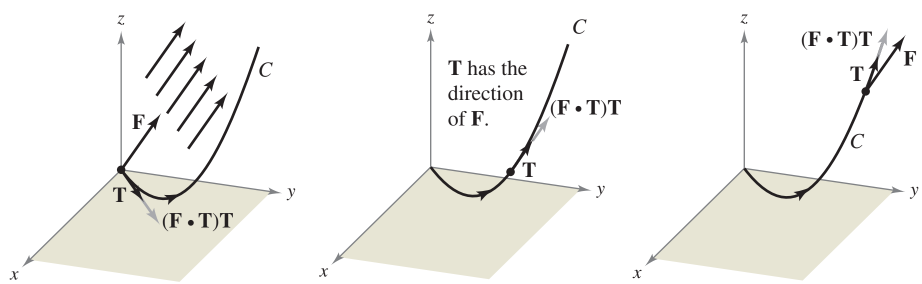
One of the most important physical applications of line integrals is that of finding the **work** done on an object moving in a force field. For example, Figure 15.12 shows an inverse square force field similar to the gravitational field of the sun. Note that the magnitude of the force along a circular path about the center is constant, whereas the magnitude of the force along a parabolic path varies from point to point.

To see how a line integral can be used to find work done in a force field $\mathbf{F}$, consider an object moving along a path C in the field, as shown in Figure 15.13. To determine the work done by the force, you need consider only that part of the force that is acting in the same direction as that in which the object is moving (or the opposite direction). This means that at each point on C, you can consider the projection $\mathbf{F} \cdot \mathbf{T}$ of the force vector $\mathbf{F}$ onto the unit tangent vector $\mathbf{T}$. On a small subarc of length Δs_i, the increment of work is

$$\begin{aligned} \Delta W_i &= (\text{force})(\text{distance}) \\ &\approx [\mathbf{F}(x_i, y_i, z_i) \cdot \mathbf{T}(x_i, y_i, z_i)]\, \Delta s_i \end{aligned}$$

where (x_i, y_i, z_i) is a point in the ith subarc. Consequently, the total work done is given by the following integral.

$$W = \int_C \mathbf{F}(x, y, z) \cdot \mathbf{T}(x, y, z)\, ds$$



At each point on C, the force in the direction of motion is $(\mathbf{F} \cdot \mathbf{T})\mathbf{T}$.

Figure 15.13

This line integral appears in other contexts and is the basis of the following definition of the **line integral of a vector field.** Note in the definition that

$$\begin{aligned} \mathbf{F} \cdot \mathbf{T}\, ds &= \mathbf{F} \cdot \frac{\mathbf{r}'(t)}{\|\mathbf{r}'(t)\|} \|\mathbf{r}'(t)\|\, dt \\ &= \mathbf{F} \cdot \mathbf{r}'(t)\, dt \\ &= \mathbf{F} \cdot d\mathbf{r}. \end{aligned}$$

Definition of Line Integral of a Vector Field

Let $\mathbf{F}$ be a continuous vector field defined on a smooth curve C given by $\mathbf{r}(t)$, $a \le t \le b$. The **line integral** of $\mathbf{F}$ on C is given by

$$\int_C \mathbf{F} \cdot d\mathbf{r} = \int_C \mathbf{F} \cdot \mathbf{T}\, ds = \int_a^b \mathbf{F}(x(t),\ y(t), z(t)) \cdot \mathbf{r}'(t)\, dt.$$

EXAMPLE 6 Work Done by a Force

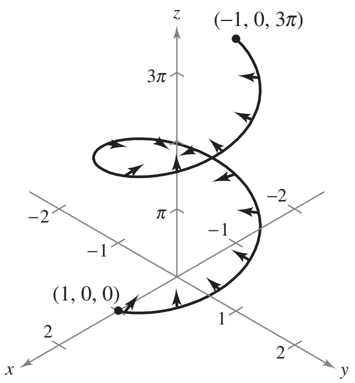


Figure 15.14

Find the work done by the force field

$$\mathbf{F}(x, y, z) = -\frac{1}{2}x\mathbf{i} - \frac{1}{2}y\mathbf{j} + \frac{1}{4}\mathbf{k} \qquad \text{Force field } \mathbf{F}$$

on a particle as it moves along the helix given by

$$\mathbf{r}(t) = \cos t\mathbf{i} + \sin t\mathbf{j} + t\mathbf{k} \qquad \text{Space curve } C$$

from the point $(1, 0, 0)$ to $(-1, 0, 3\pi)$, as shown in Figure 15.14.

Solution Because

$$\begin{aligned}\mathbf{r}(t) &= x(t)\mathbf{i} + y(t)\mathbf{j} + z(t)\mathbf{k}\\ &= \cos t\mathbf{i} + \sin t\mathbf{j} + t\mathbf{k}\end{aligned}$$

it follows that $x(t) = \cos t$, $y(t) = \sin t$, and $z(t) = t$. So, the force field can be written as

$$\mathbf{F}(x(t), y(t), z(t)) = -\frac{1}{2}\cos t\mathbf{i} - \frac{1}{2}\sin t\mathbf{j} + \frac{1}{4}\mathbf{k}.$$

To find the work done by the force field in moving a particle along the curve C, use the fact that

$$\mathbf{r}'(t) = -\sin t\mathbf{i} + \cos t\mathbf{j} + \mathbf{k}$$

and write the following.

$$\begin{aligned}W &= \int_C \mathbf{F} \cdot d\mathbf{r}\\ &= \int_a^b \mathbf{F}(x(t), y(t), z(t)) \cdot \mathbf{r}'(t)\, dt\\ &= \int_0^{3\pi} \left(-\frac{1}{2}\cos t\mathbf{i} - \frac{1}{2}\sin t\mathbf{j} + \frac{1}{4}\mathbf{k}\right) \cdot (-\sin t\mathbf{i} + \cos t\mathbf{j} + \mathbf{k})\, dt\\ &= \int_0^{3\pi} \left(\frac{1}{2}\sin t\cos t - \frac{1}{2}\sin t\cos t + \frac{1}{4}\right) dt\\ &= \int_0^{3\pi} \frac{1}{4}\, dt\\ &= \frac{1}{4}t\Big]_0^{3\pi}\\ &= \frac{3\pi}{4}\end{aligned}$$

Try It | Open Exploration

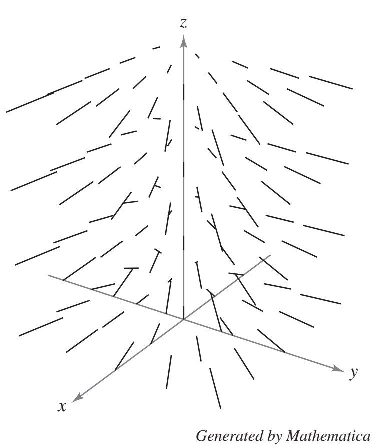


Figure 15.15

NOTE In Example 6, note that the x- and y-components of the force field end up contributing nothing to the total work. This occurs because *in this particular example* the z-component of the force field is the only portion of the force that is acting in the same (or opposite) direction in which the particle is moving (see Figure 15.15).

TECHNOLOGY The computer-generated view of the force field in Example 6 shown in Figure 15.15 indicates that each vector in the force field points toward the z-axis.

For line integrals of vector functions, the orientation of the curve C is important. If the orientation of the curve is reversed, the unit tangent vector $\mathbf{T}(t)$ is changed to $-\mathbf{T}(t)$, and you obtain

$$\int_{-C} \mathbf{F} \cdot d\mathbf{r} = -\int_C \mathbf{F} \cdot d\mathbf{r}.$$

EXAMPLE 7 Orientation and Parametrization of a Curve

Let $\mathbf{F}(x, y) = y\mathbf{i} + x^2\mathbf{j}$ and evaluate the line integral $\int_C \mathbf{F} \cdot d\mathbf{r}$ for each parabolic curve shown in Figure 15.16.

a. C_1: $\mathbf{r}_1(t) = (4 - t)\mathbf{i} + (4t - t^2)\mathbf{j}, \quad 0 \le t \le 3$

b. C_2: $\mathbf{r}_2(t) = t\mathbf{i} + (4t - t^2)\mathbf{j}, \quad 1 \le t \le 4$

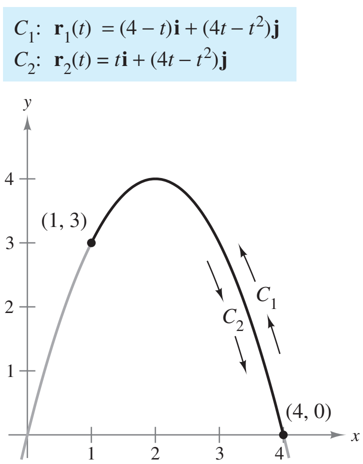


Figure 15.16

Solution

a. Because $\mathbf{r}_1'(t) = -\mathbf{i} + (4 - 2t)\mathbf{j}$ and

$$\mathbf{F}(x(t), y(t)) = (4t - t^2)\mathbf{i} + (4 - t)^2\mathbf{j}$$

the line integral is

$$\begin{aligned}
\int_{C_1} \mathbf{F} \cdot d\mathbf{r} &= \int_0^3 [(4t - t^2)\mathbf{i} + (4 - t)^2\mathbf{j}] \cdot [-\mathbf{i} + (4 - 2t)\mathbf{j}]\, dt \\
&= \int_0^3 (-4t + t^2 + 64 - 64t + 20t^2 - 2t^3)\, dt \\
&= \int_0^3 (-2t^3 + 21t^2 - 68t + 64)\, dt \\
&= \left[-\frac{t^4}{2} + 7t^3 - 34t^2 + 64t\right]_0^3 \\
&= \frac{69}{2}.
\end{aligned}$$

b. Because $\mathbf{r}_2'(t) = \mathbf{i} + (4 - 2t)\mathbf{j}$ and

$$\mathbf{F}(x(t), y(t)) = (4t - t^2)\mathbf{i} + t^2\mathbf{j}$$

the line integral is

$$\begin{aligned}
\int_{C_2} \mathbf{F} \cdot d\mathbf{r} &= \int_1^4 [(4t - t^2)\mathbf{i} + t^2\mathbf{j}] \cdot [\mathbf{i} + (4 - 2t)\mathbf{j}]\, dt \\
&= \int_1^4 (4t - t^2 + 4t^2 - 2t^3)\, dt \\
&= \int_1^4 (-2t^3 + 3t^2 + 4t)\, dt \\
&= \left[-\frac{t^4}{2} + t^3 + 2t^2\right]_1^4 \\
&= -\frac{69}{2}.
\end{aligned}$$

The answer in part (b) is the negative of that in part (a) because C_1 and C_2 represent opposite orientations of the same parabolic segment.

NOTE Although the value of the line integral in Example 7 depends on the orientation of C, it does not depend on the parametrization of C. To see this, let C_3 be represented by

$$\mathbf{r}_3 = (t + 2)\mathbf{i} + (4 - t^2)\mathbf{j}$$

where $-1 \le t \le 2$. The graph of this curve is the same parabolic segment shown in Figure 15.16. Does the value of the line integral over C_3 agree with the value over C_1 or C_2? Why or why not?

Try It

Exploration A

Line Integrals in Differential Form

A second commonly used form of line integrals is derived from the vector field notation used in the preceding section. If $\mathbf{F}$ is a vector field of the form $\mathbf{F}(x, y) = M\mathbf{i} + N\mathbf{j}$, and C is given by $\mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j}$, then $\mathbf{F} \cdot d\mathbf{r}$ is often written as $M\,dx + N\,dy$.

$$\begin{aligned}\int_C \mathbf{F} \cdot d\mathbf{r} &= \int_C \mathbf{F} \cdot \frac{d\mathbf{r}}{dt}\,dt \\ &= \int_a^b (M\mathbf{i} + N\mathbf{j}) \cdot (x'(t)\mathbf{i} + y'(t)\mathbf{j})\,dt \\ &= \int_a^b \left(M\frac{dx}{dt} + N\frac{dy}{dt}\right)dt \\ &= \int_C (M\,dx + N\,dy)\end{aligned}$$

This **differential form** can be extended to three variables. The parentheses are often omitted, as follows.

$$\int_C M\,dx + N\,dy \qquad \text{and} \qquad \int_C M\,dx + N\,dy + P\,dz$$

Notice how this differential notation is used in Example 8.

EXAMPLE 8 Evaluating a Line Integral in Differential Form

Let C be the circle of radius 3 given by

$$\mathbf{r}(t) = 3\cos t\,\mathbf{i} + 3\sin t\,\mathbf{j}, \quad 0 \le t \le 2\pi$$

as shown in Figure 15.17. Evaluate the line integral

$$\int_C y^3\,dx + (x^3 + 3xy^2)\,dy.$$

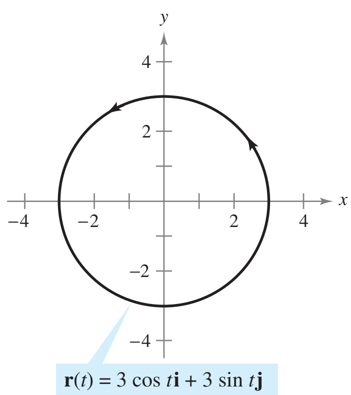


Figure 15.17

Editable Graph

Solution Because $x = 3\cos t$ and $y = 3\sin t$, you have $dx = -3\sin t\,dt$ and $dy = 3\cos t\,dt$. So, the line integral is

$$\begin{aligned}&\int_C M\,dx + N\,dy \\ &= \int_C y^3\,dx + (x^3 + 3xy^2)\,dy \\ &= \int_0^{2\pi} \left[(27\sin^3 t)(-3\sin t) + (27\cos^3 t + 81\cos t\sin^2 t)(3\cos t)\right]dt \\ &= 81\int_0^{2\pi} (\cos^4 t - \sin^4 t + 3\cos^2 t\sin^2 t)\,dt \\ &= 81\int_0^{2\pi} \left(\cos^2 t - \sin^2 t + \frac{3}{4}\sin^2 2t\right)dt \\ &= 81\int_0^{2\pi} \left[\cos 2t + \frac{3}{4}\left(\frac{1 - \cos 4t}{2}\right)\right]dt \\ &= 81\left[\frac{\sin 2t}{2} + \frac{3}{8}t - \frac{3\sin 4t}{32}\right]_0^{2\pi} \\ &= \frac{243\pi}{4}.\end{aligned}$$

Try It Exploration A

NOTE The orientation of C affects the value of the differential form of a line integral. Specifically, if $-C$ has the orientation opposite to that of C, then

$$\int_{-C} M\,dx + N\,dy = -\int_C M\,dx + N\,dy.$$

So, of the three line integral forms presented in this section, the orientation of C does not affect the form $\int_C f(x, y)\,ds$, but it does affect the vector form and the differential form.

For curves represented by $y = g(x)$, $a \le x \le b$, you can let $x = t$ and obtain the parametric form

$$x = t \quad \text{and} \quad y = g(t), \quad a \le t \le b.$$

Because $dx = dt$ for this form, you have the option of evaluating the line integral in the variable x or t. This is demonstrated in Example 9.

EXAMPLE 9 Evaluating a Line Integral in Differential Form

Evaluate

$$\int_C y\,dx + x^2\,dy$$

where C is the parabolic arc given by $y = 4x - x^2$ from $(4, 0)$ to $(1, 3)$, as shown in Figure 15.18.

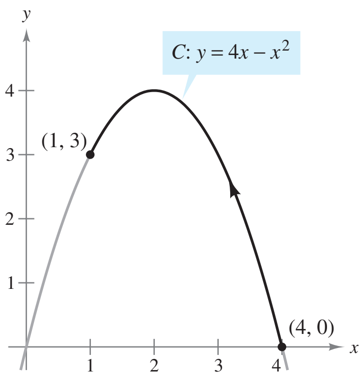


Figure 15.18

Editable Graph

Solution Rather than converting to the parameter t, you can simply retain the variable x and write

$$y = 4x - x^2 \quad \Rightarrow \quad dy = (4 - 2x)\,dx.$$

Then, in the direction from $(4, 0)$ to $(1, 3)$, the line integral is

$$\begin{aligned}\int_C y\,dx + x^2\,dy &= \int_4^1 [(4x - x^2)\,dx + x^2(4 - 2x)\,dx] \\ &= \int_4^1 (4x + 3x^2 - 2x^3)\,dx \\ &= \left[2x^2 + x^3 - \frac{x^4}{2}\right]_4^1 = \frac{69}{2}.\end{aligned}$$

See Example 7.

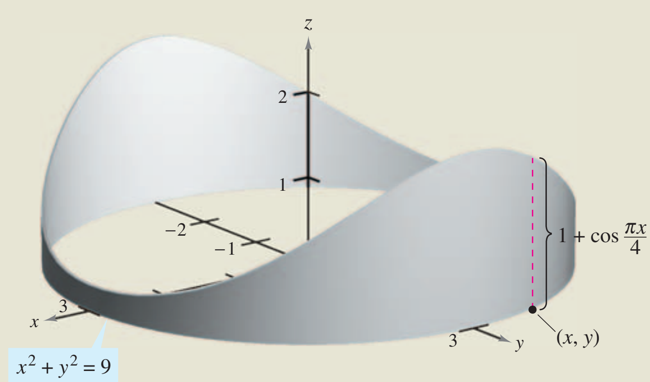
Try It Exploration A

EXPLORATION

Finding Lateral Surface Area The figure below shows a piece of tin that has been cut from a circular cylinder. The base of the circular cylinder is modeled by $x^2 + y^2 = 9$. At any point (x, y) on the base, the height of the object is given by

$$f(x, y) = 1 + \cos\frac{\pi x}{4}.$$

Explain how to use a line integral to find the surface area of the piece of tin.



Rotatable Graph

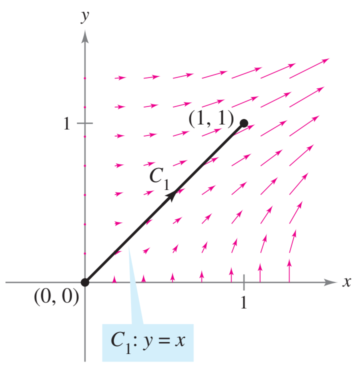
Section 15.3 Conservative Vector Fields and Independence of Path

- **Understand and use the Fundamental Theorem of Line Integrals.**
- **Understand the concept of independence of path.**
- **Understand the concept of conservation of energy.**

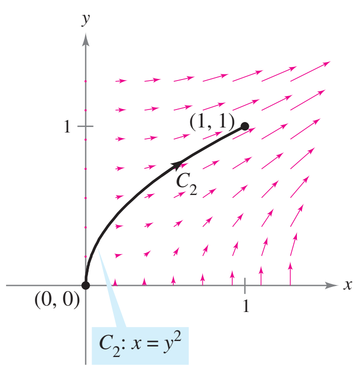
Fundamental Theorem of Line Integrals

The discussion in the preceding section pointed out that in a gravitational field the work done by gravity on an object moving between two points in the field is independent of the path taken by the object. In this section, you will study an important generalization of this result—it is called the **Fundamental Theorem of Line Integrals.**

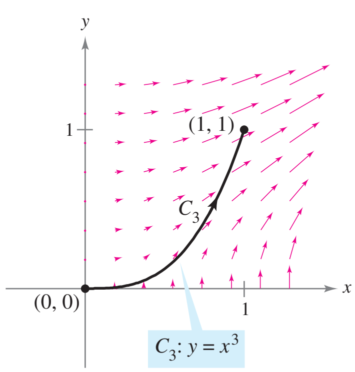
To begin, an example is presented in which the line integral of a *conservative vector field* is evaluated over three different paths.



(a)



(b)



(c)

Figure 15.19

EXAMPLE 1 Line Integral of a Conservative Vector Field

Find the work done by the force field

$$\mathbf{F}(x, y) = \frac{1}{2}xy\mathbf{i} + \frac{1}{4}x^2\mathbf{j}$$

on a particle that moves from $(0, 0)$ to $(1, 1)$ along each path, as shown in Figure 15.19.

a. C_1: $y = x$ **b.** C_2: $x = y^2$ **c.** C_3: $y = x^3$

Solution

a. Let $\mathbf{r}(t) = t\mathbf{i} + t\mathbf{j}$ for $0 \le t \le 1$, so that

$$d\mathbf{r} = (\mathbf{i} + \mathbf{j})\,dt \quad \text{and} \quad \mathbf{F}(x, y) = \frac{1}{2}t^2\mathbf{i} + \frac{1}{4}t^2\mathbf{j}.$$

Then, the work done is

$$W = \int_{C_1} \mathbf{F} \cdot d\mathbf{r} = \int_0^1 \frac{3}{4}t^2\,dt = \frac{1}{4}t^3\Big]_0^1 = \frac{1}{4}.$$

b. Let $\mathbf{r}(t) = t\mathbf{i} + \sqrt{t}\mathbf{j}$ for $0 \le t \le 1$, so that

$$d\mathbf{r} = \left(\mathbf{i} + \frac{1}{2\sqrt{t}}\mathbf{j}\right)dt \quad \text{and} \quad \mathbf{F}(x, y) = \frac{1}{2}t^{3/2}\mathbf{i} + \frac{1}{4}t^2\mathbf{j}.$$

Then, the work done is

$$W = \int_{C_2} \mathbf{F} \cdot d\mathbf{r} = \int_0^1 \frac{5}{8}t^{3/2}\,dt = \frac{1}{4}t^{5/2}\Big]_0^1 = \frac{1}{4}.$$

c. Let $\mathbf{r}(t) = \frac{1}{2}t\mathbf{i} + \frac{1}{8}t^3\mathbf{j}$ for $0 \le t \le 2$, so that

$$d\mathbf{r} = \left(\frac{1}{2}\mathbf{i} + \frac{3}{8}t^2\mathbf{j}\right)dt \quad \text{and} \quad \mathbf{F}(x, y) = \frac{1}{32}t^4\mathbf{i} + \frac{1}{16}t^2\mathbf{j}.$$

Then, the work done is

$$W = \int_{C_3} \mathbf{F} \cdot d\mathbf{r} = \int_0^2 \frac{5}{128}t^4\,dt = \frac{1}{128}t^5\Big]_0^2 = \frac{1}{4}.$$

So, the work done by a conservative vector field is the same for all paths.

Try It **Exploration A**

In Example 1, note that the vector field $\mathbf{F}(x, y) = \frac{1}{2}xy\mathbf{i} + \frac{1}{4}x^2\mathbf{j}$ is conservative because $\mathbf{F}(x, y) = \nabla f(x, y)$, where $f(x, y) = \frac{1}{4}x^2y$. In such cases, the following theorem states that the value of $\int_C \mathbf{F} \cdot d\mathbf{r}$ is given by

$$\begin{aligned}\int_C \mathbf{F} \cdot d\mathbf{r} &= f(x(1), y(1)) - f(x(0), y(0))\\ &= \frac{1}{4} - 0\\ &= \frac{1}{4}.\end{aligned}$$

THEOREM 15.5 Fundamental Theorem of Line Integrals

Let C be a piecewise smooth curve lying in an open region R and given by

$$\mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j}, \quad a \leq t \leq b.$$

If $\mathbf{F}(x, y) = M\mathbf{i} + N\mathbf{j}$ is conservative in R, and M and N are continuous in R, then

$$\int_C \mathbf{F} \cdot d\mathbf{r} = \int_C \nabla f \cdot d\mathbf{r} = f(x(b), y(b)) - f(x(a), y(a))$$

where f is a potential function of $\mathbf{F}$. That is, $\mathbf{F}(x, y) = \nabla f(x, y)$.

NOTE Notice how the Fundamental Theorem of Line Integrals is similar to the Fundamental Theorem of Calculus (Section 4.4), which states that

$$\int_a^b f(x)\,dx = F(b) - F(a)$$

where $F'(x) = f(x)$.

Proof A proof is provided only for a smooth curve. For piecewise smooth curves, the procedure is carried out separately on each smooth portion. Because $\mathbf{F}(x, y) = \nabla f(x, y) = f_x(x, y)\mathbf{i} + f_y(x, y)\mathbf{j}$, it follows that

$$\begin{aligned}\int_C \mathbf{F} \cdot d\mathbf{r} &= \int_a^b \mathbf{F} \cdot \frac{d\mathbf{r}}{dt}\,dt\\ &= \int_a^b \left[f_x(x, y)\frac{dx}{dt} + f_y(x, y)\frac{dy}{dt}\right] dt\end{aligned}$$

and, by the Chain Rule (Theorem 13.6), you have

$$\begin{aligned}\int_C \mathbf{F} \cdot d\mathbf{r} &= \int_a^b \frac{d}{dt}[f(x(t), y(t))]\,dt\\ &= f(x(b), y(b)) - f(x(a), y(a)).\end{aligned}$$

The last step is an application of the Fundamental Theorem of Calculus.

In space, the Fundamental Theorem of Line Integrals takes the following form. Let C be a piecewise smooth curve lying in an open region Q and given by

$$\mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j} + z(t)\mathbf{k}, \quad a \leq t \leq b.$$

If $\mathbf{F}(x, y, z) = M\mathbf{i} + N\mathbf{j} + P\mathbf{k}$ is conservative and M, N, and P are continuous, then

$$\begin{aligned}\int_C \mathbf{F} \cdot d\mathbf{r} &= \int_C \nabla f \cdot d\mathbf{r}\\ &= f(x(b), y(b), z(b)) - f(x(a), y(a), z(a))\end{aligned}$$

where $\mathbf{F}(x, y, z) = \nabla f(x, y, z)$.

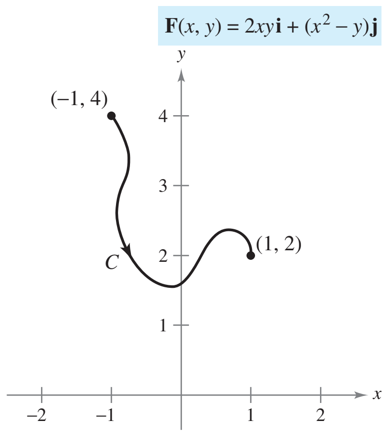
The Fundamental Theorem of Line Integrals states that if the vector field $\mathbf{F}$ is conservative, then the line integral between any two points is simply the difference in the values of the *potential* function f at these points.

EXAMPLE 2 Using the Fundamental Theorem of Line Integrals

Evaluate $\int_C \mathbf{F} \cdot d\mathbf{r}$, where C is a piecewise smooth curve from $(-1, 4)$ to $(1, 2)$ and

$$\mathbf{F}(x, y) = 2xy\mathbf{i} + (x^2 - y)\mathbf{j}$$

as shown in Figure 15.20.



Using the Fundamental Theorem of Line Integrals, $\int_C \mathbf{F} \cdot d\mathbf{r}$.
Figure 15.20

Solution From Example 6 in Section 15.1, you know that $\mathbf{F}$ is the gradient of f where

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

Consequently, $\mathbf{F}$ is conservative, and by the Fundamental Theorem of Line Integrals, it follows that

$$\begin{aligned}\int_C \mathbf{F} \cdot d\mathbf{r} &= f(1, 2) - f(-1, 4) \\ &= \left[1^2(2) - \frac{2^2}{2}\right] - \left[(-1)^2(4) - \frac{4^2}{2}\right] \\ &= 4.\end{aligned}$$

Note that it is unnecessary to include a constant K as part of f, because it is canceled by subtraction.

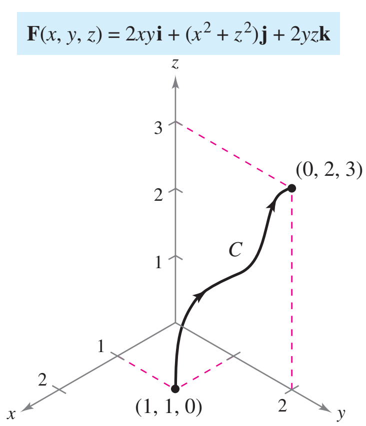
Try It   Exploration A

EXAMPLE 3 Using the Fundamental Theorem of Line Integrals

Evaluate $\int_C \mathbf{F} \cdot d\mathbf{r}$, where C is a piecewise smooth curve from $(1, 1, 0)$ to $(0, 2, 3)$ and

$$\mathbf{F}(x, y, z) = 2xy\mathbf{i} + (x^2 + z^2)\mathbf{j} + 2yz\mathbf{k}$$

as shown in Figure 15.21.



Using the Fundamental Theorem of Line Integrals, $\int_C \mathbf{F} \cdot d\mathbf{r}$.
Figure 15.21

Solution From Example 8 in Section 15.1, you know that $\mathbf{F}$ is the gradient of f where $f(x, y, z) = x^2y + yz^2 + K$. Consequently, $\mathbf{F}$ is conservative, and by the Fundamental Theorem of Line Integrals, it follows that

$$\begin{aligned}\int_C \mathbf{F} \cdot d\mathbf{r} &= f(0, 2, 3) - f(1, 1, 0) \\ &= [(0)^2(2) + (2)(3)^2] - [(1)^2(1) + (1)(0)^2] \\ &= 17.\end{aligned}$$

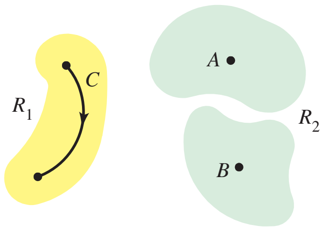
Try It   Exploration A

In Examples 2 and 3, be sure you see that the value of the line integral is the same for any smooth curve C that has the given initial and terminal points. For instance, in Example 3, try evaluating the line integral for the curve given by

$$\mathbf{r}(t) = (1 - t)\mathbf{i} + (1 + t)\mathbf{j} + 3t\mathbf{k}.$$

You should obtain

$$\begin{aligned}\int_C \mathbf{F} \cdot d\mathbf{r} &= \int_0^1 (30t^2 + 16t - 1)\, dt \\ &= 17.\end{aligned}$$



R_1 is connected. R_2 is not connected.

Figure 15.22

Independence of Path

From the Fundamental Theorem of Line Integrals it is clear that if $\mathbf{F}$ is continuous and conservative in an open region R, the value of $\int_C \mathbf{F} \cdot d\mathbf{r}$ is the same for every piecewise smooth curve C from one fixed point in R to another fixed point in R. This result is described by saying that the line integral $\int_C \mathbf{F} \cdot d\mathbf{r}$ is **independent of path** in the region R.

A region in the plane (or in space) is **connected** if any two points in the region can be joined by a piecewise smooth curve lying entirely within the region, as shown in Figure 15.22. In open regions that are *connected*, the path independence of $\int_C \mathbf{F} \cdot d\mathbf{r}$ is equivalent to the condition that $\mathbf{F}$ is conservative.

THEOREM 15.6 Independence of Path and Conservative Vector Fields

If $\mathbf{F}$ is continuous on an open connected region, then the line integral

$$\int_C \mathbf{F} \cdot d\mathbf{r}$$

is independent of path if and only if $\mathbf{F}$ is conservative.

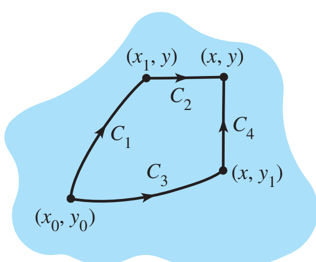


Figure 15.23

Proof If $\mathbf{F}$ is conservative, then, by the Fundamental Theorem of Line Integrals, the line integral is independent of path. Now establish the converse for a plane region R. Let $\mathbf{F}(x, y) = M\mathbf{i} + N\mathbf{j}$, and let (x_0, y_0) be a fixed point in R. If (x, y) is any point in R, choose a piecewise smooth curve C running from (x_0, y_0) to (x, y), and define f by

$$f(x, y) = \int_C \mathbf{F} \cdot d\mathbf{r} = \int_C M\,dx + N\,dy.$$

The existence of C in R is guaranteed by the fact that R is connected. You can show that f is a potential function of $\mathbf{F}$ by considering two different paths between (x_0, y_0) and (x, y). For the *first* path, choose (x_1, y) in R such that $x \neq x_1$. This is possible because R is open. Then choose C_1 and C_2, as shown in Figure 15.23. Using the independence of path, it follows that

$$\begin{aligned} f(x, y) &= \int_C M\,dx + N\,dy \\ &= \int_{C_1} M\,dx + N\,dy + \int_{C_2} M\,dx + N\,dy. \end{aligned}$$

Because the first integral does not depend on x, and because $dy = 0$ in the second integral, you have

$$f(x, y) = g(y) + \int_{C_2} M\,dx$$

and it follows that the partial derivative of f with respect to x is $f_x(x, y) = M$. For the *second* path, choose a point (x, y_1). Using reasoning similar to that used for the first path, you can conclude that $f_y(x, y) = N$. Therefore,

$$\begin{aligned} \nabla f(x, y) &= f_x(x, y)\mathbf{i} + f_y(x, y)\mathbf{j} \\ &= M\mathbf{i} + N\mathbf{j} \\ &= \mathbf{F}(x, y) \end{aligned}$$

and it follows that $\mathbf{F}$ is conservative.

EXAMPLE 4 Finding Work in a Conservative Force Field

For the force field given by

$$\mathbf{F}(x, y, z) = e^x \cos y\mathbf{i} - e^x \sin y\mathbf{j} + 2\mathbf{k}$$

show that $\int_C \mathbf{F} \cdot d\mathbf{r}$ is independent of path, and calculate the work done by $\mathbf{F}$ on an object moving along a curve C from $(0, \pi/2, 1)$ to $(1, \pi, 3)$.

Solution Writing the force field in the form $\mathbf{F}(x, y, z) = M\mathbf{i} + N\mathbf{j} + P\mathbf{k}$, you have $M = e^x \cos y$, $N = -e^x \sin y$, and $P = 2$, and it follows that

$$\frac{\partial P}{\partial y} = 0 = \frac{\partial N}{\partial z}$$
$$\frac{\partial P}{\partial x} = 0 = \frac{\partial M}{\partial z}$$
$$\frac{\partial N}{\partial x} = -e^x \sin y = \frac{\partial M}{\partial y}.$$

So, $\mathbf{F}$ is conservative. If f is a potential function of $\mathbf{F}$, then

$$f_x(x, y, z) = e^x \cos y$$
$$f_y(x, y, z) = -e^x \sin y$$
$$f_z(x, y, z) = 2.$$

By integrating with respect to x, y, and z separately, you obtain

$$f(x, y, z) = \int f_x(x, y, z)\,dx = \int e^x \cos y\,dx = e^x \cos y + g(y, z)$$
$$f(x, y, z) = \int f_y(x, y, z)\,dy = \int -e^x \sin y\,dy = e^x \cos y + h(x, z)$$
$$f(x, y, z) = \int f_z(x, y, z)\,dz = \int 2\,dz = 2z + k(x, y).$$

By comparing these three versions of $f(x, y, z)$, you can conclude that

$$f(x, y, z) = e^x \cos y + 2z + K.$$

Therefore, the work done by $\mathbf{F}$ along *any* curve C from $(0, \pi/2, 1)$ to $(1, \pi, 3)$ is

$$\begin{aligned} W &= \int_C \mathbf{F} \cdot d\mathbf{r} \\ &= \Big[e^x \cos y + 2z\Big]_{(0, \pi/2, 1)}^{(1, \pi, 3)} \\ &= (-e + 6) - (0 + 2) \\ &= 4 - e. \end{aligned}$$

Try It **Exploration A**

How much work would be done if the object in Example 4 moved from the point $(0, \pi/2, 1)$ to $(1, \pi, 3)$ and then back to the starting point $(0, \pi/2, 1)$? The Fundamental Theorem of Line Integrals states that there is zero work done. Remember that, by definition, work can be negative. So, by the time the object gets back to its starting point, the amount of work that registers positively is canceled out by the amount of work that registers negatively.

A curve C given by $\mathbf{r}(t)$ for $a \le t \le b$ is **closed** if $\mathbf{r}(a) = \mathbf{r}(b)$. By the Fundamental Theorem of Line Integrals, you can conclude that if $\mathbf{F}$ is continuous and conservative on an open region R, then the line integral over every closed curve C is 0.

NOTE Theorem 15.7 gives you options for evaluating a line integral involving a conservative vector field. You can use a potential function, or it might be more convenient to choose a particularly simple path, such as a straight line.

THEOREM 15.7 Equivalent Conditions

Let $\mathbf{F}(x, y, z) = M\mathbf{i} + N\mathbf{j} + P\mathbf{k}$ have continuous first partial derivatives in an open connected region R, and let C be a piecewise smooth curve in R. The following conditions are equivalent.

1. $\mathbf{F}$ is conservative. That is, $\mathbf{F} = \nabla f$ for some function f.
2. $\displaystyle\int_C \mathbf{F} \cdot d\mathbf{r}$ is independent of path.
3. $\displaystyle\int_C \mathbf{F} \cdot d\mathbf{r} = 0$ for every *closed* curve C in R.

EXAMPLE 5 Evaluating a Line Integral

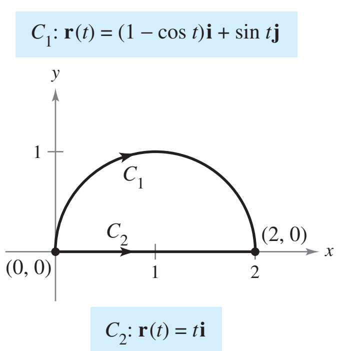


Figure 15.24

Evaluate $\displaystyle\int_{C_1} \mathbf{F} \cdot d\mathbf{r}$, where

$$\mathbf{F}(x, y) = (y^3 + 1)\mathbf{i} + (3xy^2 + 1)\mathbf{j}$$

and C_1 is the semicircular path from $(0, 0)$ to $(2, 0)$, as shown in Figure 15.24.

Solution You have the following three options.

a. You can use the method presented in the preceding section to evaluate the line integral along the *given curve*. To do this, you can use the parametrization $\mathbf{r}(t) = (1 - \cos t)\mathbf{i} + \sin t\mathbf{j}$, where $0 \le t \le \pi$. For this parametrization, it follows that $d\mathbf{r} = \mathbf{r}'(t)\,dt = (\sin t\mathbf{i} + \cos t\mathbf{j})\,dt$, and

$$\int_{C_1} \mathbf{F} \cdot d\mathbf{r} = \int_0^{\pi} (\sin t + \sin^4 t + \cos t + 3\sin^2 t \cos t - 3\sin^2 t \cos^2 t)\,dt.$$

This integral should dampen your enthusiasm for this option.

b. You can try to find a *potential function* and evaluate the line integral by the Fundamental Theorem of Line Integrals. Using the technique demonstrated in Example 4, you can find the potential function to be $f(x, y) = xy^3 + x + y + K$, and, by the Fundamental Theorem,

$$W = \int_{C_1} \mathbf{F} \cdot d\mathbf{r} = f(2, 0) - f(0, 0) = 2.$$

c. Knowing that $\mathbf{F}$ is conservative, you have a third option. Because the value of the line integral is independent of path, you can replace the semicircular path with a *simpler path*. Suppose you choose the straight-line path C_2 from $(0, 0)$ to $(2, 0)$. Then, $\mathbf{r}(t) = t\mathbf{i}$, where $0 \le t \le 2$. So, $d\mathbf{r} = \mathbf{i}\,dt$ and $\mathbf{F}(x, y) = (y^3 + 1)\mathbf{i} + (3xy^2 + 1)\mathbf{j} = \mathbf{i} + \mathbf{j}$, so that

$$\int_{C_1} \mathbf{F} \cdot d\mathbf{r} = \int_{C_2} \mathbf{F} \cdot d\mathbf{r} = \int_0^2 1\,dt = t\Big]_0^2 = 2.$$

Of the three options, obviously the third one is the easiest.

Try It | Exploration A | Open Exploration

MICHAEL FARADAY (1791–1867)

Several philosophers of science have considered Faraday's Law of Conservation of Energy to be the greatest generalization ever conceived by humankind. Many physicists have contributed to our knowledge of this law. Two early and influential ones were James Prescott Joule (1818–1889) and Hermann Ludwig Helmholtz (1821–1894).

Conservation of Energy

In 1840, the English physicist Michael Faraday wrote, "Nowhere is there a pure creation or production of power without a corresponding exhaustion of something to supply it." This statement represents the first formulation of one of the most important laws of physics—**the Law of Conservation of Energy.** In modern terminology, the law is stated as follows: *In a conservative force field, the sum of the potential and kinetic energies of an object remains constant from point to point.*

You can use the Fundamental Theorem of Line Integrals to derive this law. From physics, the **kinetic energy** of a particle of mass m and speed v is $k = \frac{1}{2}mv^2$. The **potential energy** p of a particle at point (x, y, z) in a conservative vector field $\mathbf{F}$ is defined as $p(x, y, z) = -f(x, y, z)$, where f is the potential function for $\mathbf{F}$. Consequently, the work done by $\mathbf{F}$ along a smooth curve C from A to B is

$$\begin{aligned} W = \int_C \mathbf{F} \cdot d\mathbf{r} &= f(x, y, z)\Big]_A^B \\ &= -p(x, y, z)\Big]_A^B \\ &= p(A) - p(B) \end{aligned}$$

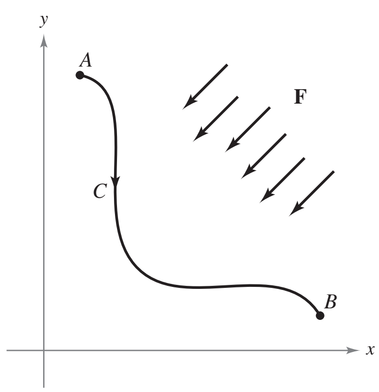
as shown in Figure 15.25. In other words, work W is equal to the difference in the potential energies of A and B. Now, suppose that $\mathbf{r}(t)$ is the position vector for a particle moving along C from $A = \mathbf{r}(a)$ to $B = \mathbf{r}(b)$. At any time t, the particle's velocity, acceleration, and speed are $\mathbf{v}(t) = \mathbf{r}'(t)$, $\mathbf{a}(t) = \mathbf{r}''(t)$, and $v(t) = \|\mathbf{v}(t)\|$, respectively. So, by Newton's Second Law of Motion, $\mathbf{F} = m\mathbf{a}(t) = m(\mathbf{v}'(t))$, and the work done by $\mathbf{F}$ is

$$\begin{aligned} W = \int_C \mathbf{F} \cdot d\mathbf{r} &= \int_a^b \mathbf{F} \cdot \mathbf{r}'(t)\, dt \\ &= \int_a^b \mathbf{F} \cdot \mathbf{v}(t)\, dt = \int_a^b [m\mathbf{v}'(t)] \cdot \mathbf{v}(t)\, dt \\ &= \int_a^b m[\mathbf{v}'(t) \cdot \mathbf{v}(t)]\, dt \\ &= \frac{m}{2}\int_a^b \frac{d}{dt}[\mathbf{v}(t) \cdot \mathbf{v}(t)]\, dt \\ &= \frac{m}{2}\int_a^b \frac{d}{dt}[\|\mathbf{v}(t)\|^2]\, dt \\ &= \frac{m}{2}\Big[\|\mathbf{v}(t)\|^2\Big]_a^b \\ &= \frac{m}{2}\Big[[v(t)]^2\Big]_a^b \\ &= \frac{1}{2}m[v(b)]^2 - \frac{1}{2}m[v(a)]^2 \\ &= k(B) - k(A). \end{aligned}$$

Equating these two results for W produces

$$\begin{aligned} p(A) - p(B) &= k(B) - k(A) \\ p(A) + k(A) &= p(B) + k(B) \end{aligned}$$

which implies that the sum of the potential and kinetic energies remains constant from point to point.



The work done by $\mathbf{F}$ along C is
$W = \int_C \mathbf{F} \cdot d\mathbf{r} = p(A) - p(B).$

Figure 15.25

Section 15.4 Green's Theorem

- **Use Green's Theorem to evaluate a line integral.**
- **Use alternative forms of Green's Theorem.**

Green's Theorem

In this section, you will study **Green's Theorem,** named after the English mathematician George Green (1793–1841). This theorem states that the value of a double integral over a *simply connected* plane region R is determined by the value of a line integral around the boundary of R.

A curve C given by $\mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j}$, where $a \le t \le b$, is **simple** if it does not cross itself—that is, $\mathbf{r}(c) \ne \mathbf{r}(d)$ for all c and d in the open interval (a, b). A plane region R is **simply connected** if its boundary consists of *one* simple closed curve, as shown in Figure 15.26.

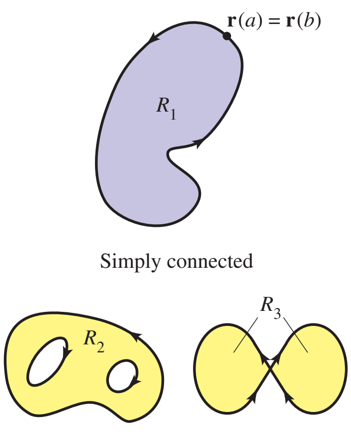


Figure 15.26

THEOREM 15.8 Green's Theorem

Let R be a simply connected region with a piecewise smooth boundary C, oriented counterclockwise (that is, C is traversed *once* so that the region R always lies to the *left*). If M and N have continuous partial derivatives in an open region containing R, then

$$\int_C M\,dx + N\,dy = \int_R\!\!\int \left(\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y}\right) dA.$$

Proof A proof is given only for a region that is both vertically simple and horizontally simple, as shown in Figure 15.27.

$$\begin{aligned}\int_C M\,dx &= \int_{C_1} M\,dx + \int_{C_2} M\,dx\\ &= \int_a^b M(x, f_1(x))\,dx + \int_b^a M(x, f_2(x))\,dx\\ &= \int_a^b [M(x, f_1(x)) - M(x, f_2(x))]\,dx\end{aligned}$$

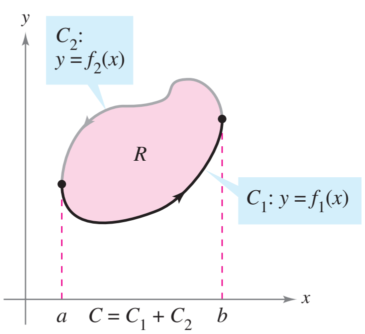
On the other hand,

$$\begin{aligned}\int_R\!\!\int \frac{\partial M}{\partial y}\,dA &= \int_a^b \int_{f_1(x)}^{f_2(x)} \frac{\partial M}{\partial y}\,dy\,dx\\ &= \int_a^b M(x, y)\Big]_{f_1(x)}^{f_2(x)}\,dx\\ &= \int_a^b [M(x, f_2(x)) - M(x, f_1(x))]\,dx.\end{aligned}$$

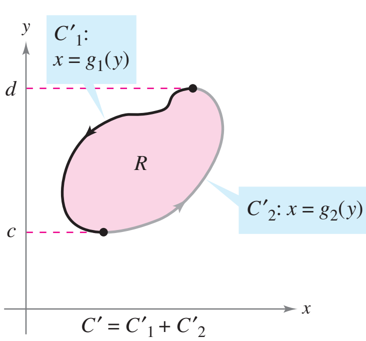
Consequently,

$$\int_C M\,dx = -\int_R\!\!\int \frac{\partial M}{\partial y}\,dA.$$

Similarly, you can use $g_1(y)$ and $g_2(y)$ to show that $\int_C N\,dy = \int_R\!\int \partial N/\partial x\,dA$. By adding the integrals $\int_C M\,dx$ and $\int_C N\,dy$, you obtain the conclusion stated in the theorem.



R is vertically simple.



R is horizontally simple.

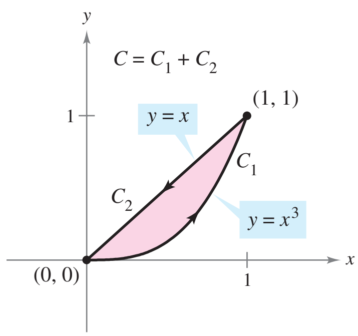
Figure 15.27

EXAMPLE 1 Using Green's Theorem

Use Green's Theorem to evaluate the line integral

$$\int_C y^3\,dx + (x^3 + 3xy^2)\,dy$$

where C is the path from $(0, 0)$ to $(1, 1)$ along the graph of $y = x^3$ and from $(1, 1)$ to $(0, 0)$ along the graph of $y = x$, as shown in Figure 15.28.



C is simple and closed, and the region R always lies to the left of C.
Figure 15.28

Solution Because $M = y^3$ and $N = x^3 + 3xy^2$, it follows that

$$\frac{\partial N}{\partial x} = 3x^2 + 3y^2 \quad \text{and} \quad \frac{\partial M}{\partial y} = 3y^2.$$

Applying Green's Theorem, you then have

$$\begin{aligned}
\int_C y^3\,dx + (x^3 + 3xy^2)\,dy &= \int_R\!\!\int \left(\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y}\right) dA \\
&= \int_0^1 \int_{x^3}^{x} [(3x^2 + 3y^2) - 3y^2]\,dy\,dx \\
&= \int_0^1 \int_{x^3}^{x} 3x^2\,dy\,dx \\
&= \int_0^1 3x^2y\Big]_{x^3}^{x}\,dx \\
&= \int_0^1 (3x^3 - 3x^5)\,dx \\
&= \left[\frac{3x^4}{4} - \frac{x^6}{2}\right]_0^1 \\
&= \frac{1}{4}.
\end{aligned}$$

Try It **Exploration A**

Green's Theorem cannot be applied to every line integral. Among other restrictions stated in Theorem 15.8, the curve C must be simple and closed. When Green's Theorem does apply, however, it can save time. To see this, try using the techniques described in Section 15.2 to evaluate the line integral in Example 1. To do this, you would need to write the line integral as

$$\int_C y^3\,dx + (x^3 + 3xy^2)\,dy =$$
$$\int_{C_1} y^3\,dx + (x^3 + 3xy^2)\,dy + \int_{C_2} y^3\,dx + (x^3 + 3xy^2)\,dy$$

where C_1 is the cubic path given by

$$\mathbf{r}(t) = t\mathbf{i} + t^3\mathbf{j}$$

from $t = 0$ to $t = 1$, and C_2 is the line segment given by

$$\mathbf{r}(t) = (1 - t)\mathbf{i} + (1 - t)\mathbf{j}$$

from $t = 0$ to $t = 1$.

GEORGE GREEN (1793–1841)

Green, a self-educated miller's son, first published the theorem that bears his name in 1828 in an essay on electricity and magnetism. At that time there was almost no mathematical theory to explain electrical phenomena. "Considering how desirable it was that a power of universal agency, like electricity, should, as far as possible, be submitted to calculation, . . . I was induced to try whether it would be possible to discover any general relations existing between this function and the quantities of electricity in the bodies producing it."

EXAMPLE 2 Using Green's Theorem to Calculate Work

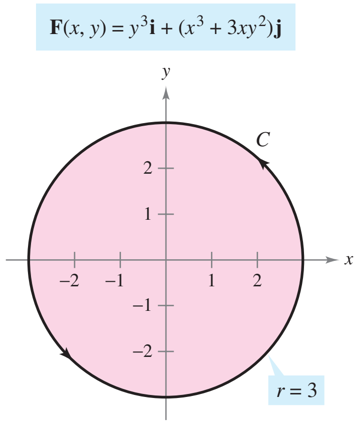


Figure 15.29

While subject to the force

$$\mathbf{F}(x, y) = y^3\mathbf{i} + (x^3 + 3xy^2)\mathbf{j}$$

a particle travels once around the circle of radius 3 shown in Figure 15.29. Use Green's Theorem to find the work done by **F**.

Solution From Example 1, you know by Green's Theorem that

$$\int_C y^3\,dx + (x^3 + 3xy^2)\,dy = \int_R\!\!\int 3x^2\,dA.$$

In polar coordinates, using $x = r\cos\theta$ and $dA = r\,dr\,d\theta$, the work done is

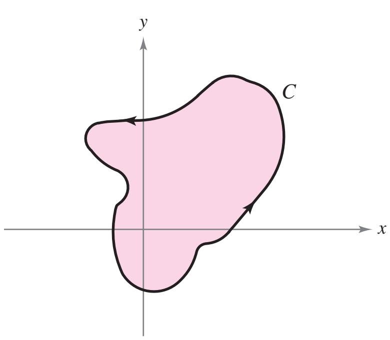
$$\begin{aligned}
W = \int_R\!\!\int 3x^2\,dA &= \int_0^{2\pi}\!\!\int_0^3 3(r\cos\theta)^2\,r\,dr\,d\theta \\
&= 3\int_0^{2\pi}\!\!\int_0^3 r^3\cos^2\theta\,dr\,d\theta \\
&= 3\int_0^{2\pi} \frac{r^4}{4}\cos^2\theta\Big]_0^3\,d\theta \\
&= 3\int_0^{2\pi} \frac{81}{4}\cos^2\theta\,d\theta \\
&= \frac{243}{8}\int_0^{2\pi}(1 + \cos 2\theta)\,d\theta \\
&= \frac{243}{8}\left[\theta + \frac{\sin 2\theta}{2}\right]_0^{2\pi} \\
&= \frac{243\pi}{4}.
\end{aligned}$$

Try It | Exploration A | Exploration B

When evaluating line integrals over closed curves, remember that for conservative vector fields (those for which $\partial N/\partial x = \partial M/\partial y$), the value of the line integral is 0. This is easily seen from the statement of Green's Theorem:

$$\int_C M\,dx + N\,dy = \int_R\!\!\int \left(\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y}\right) dA = 0.$$

EXAMPLE 3 Green's Theorem and Conservative Vector Fields



C is closed.
Figure 15.30

Evaluate the line integral

$$\int_C y^3\,dx + 3xy^2\,dy$$

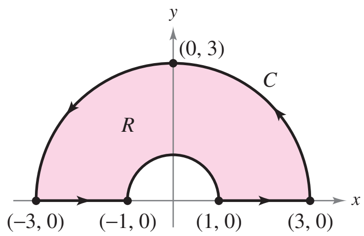
where C is the path shown in Figure 15.30.

Solution From this line integral, $M = y^3$ and $N = 3xy^2$. So, $\partial N/\partial x = 3y^2$ and $\partial M/\partial y = 3y^2$. This implies that the vector field $\mathbf{F} = M\mathbf{i} + N\mathbf{j}$ is conservative, and because C is closed, you can conclude that

$$\int_C y^3\,dx + 3xy^2\,dy = 0.$$

Try It | Exploration A

EXAMPLE 4 Using Green's Theorem for a Piecewise Smooth Curve



C is piecewise smooth.
Figure 15.31

Evaluate

$$\int_C (\arctan x + y^2)\,dx + (e^y - x^2)\,dy$$

where C is the path enclosing the annular region shown in Figure 15.31.

Solution In polar coordinates, R is given by $1 \le r \le 3$ for $0 \le \theta \le \pi$. Moreover,

$$\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y} = -2x - 2y = -2(r\cos\theta + r\sin\theta).$$

So, by Green's Theorem,

$$\begin{aligned}
\int_C (\arctan x + y^2)\,dx + (e^y - x^2)\,dy &= \int_R\!\int -2(x+y)\,dA \\
&= \int_0^{\pi}\int_1^3 -2r(\cos\theta + \sin\theta)r\,dr\,d\theta \\
&= \int_0^{\pi} -2(\cos\theta + \sin\theta)\frac{r^3}{3}\Big]_1^3\,d\theta \\
&= \int_0^{\pi}\left(-\frac{52}{3}\right)(\cos\theta + \sin\theta)\,d\theta \\
&= -\frac{52}{3}\Big[\sin\theta - \cos\theta\Big]_0^{\pi} \\
&= -\frac{104}{3}.
\end{aligned}$$

Try It | Exploration A | Open Exploration

In Examples 1, 2, and 4, Green's Theorem was used to evaluate line integrals as double integrals. You can also use the theorem to evaluate double integrals as line integrals. One useful application occurs when $\partial N/\partial x - \partial M/\partial y = 1$.

$$\begin{aligned}
\int_C M\,dx + N\,dy &= \int_R\!\int \left(\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y}\right) dA \\
&= \int_R\!\int 1\,dA \qquad \frac{\partial N}{\partial x} - \frac{\partial M}{\partial y} = 1 \\
&= \text{area of region } R
\end{aligned}$$

Among the many choices for M and N satisfying the stated condition, the choice of $M = -y/2$ and $N = x/2$ produces the following line integral for the area of region R.

THEOREM 15.9 Line Integral for Area

If R is a plane region bounded by a piecewise smooth simple closed curve C, oriented counterclockwise, then the area of R is given by

$$A = \frac{1}{2}\int_C x\,dy - y\,dx.$$

EXAMPLE 5 Finding Area by a Line Integral

Use a line integral to find the area of the ellipse

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1.$$

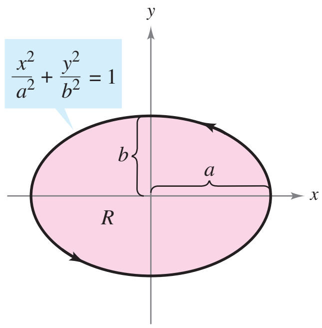


Figure 15.32

Solution Using Figure 15.32, you can induce a counterclockwise orientation to the elliptical path by letting

$$x = a\cos t \quad \text{and} \quad y = b\sin t, \quad 0 \le t \le 2\pi.$$

So, the area is

$$\begin{aligned} A = \frac{1}{2}\int_C x\,dy - y\,dx &= \frac{1}{2}\int_0^{2\pi} [(a\cos t)(b\cos t)\,dt - (b\sin t)(-a\sin t)\,dt] \\ &= \frac{ab}{2}\int_0^{2\pi} (\cos^2 t + \sin^2 t)\,dt \\ &= \frac{ab}{2}\Big[t\Big]_0^{2\pi} \\ &= \pi ab. \end{aligned}$$

Try It **Exploration A**

Green's Theorem can be extended to cover some regions that are not simply connected. This is demonstrated in the next example.

EXAMPLE 6 Green's Theorem Extended to a Region with a Hole

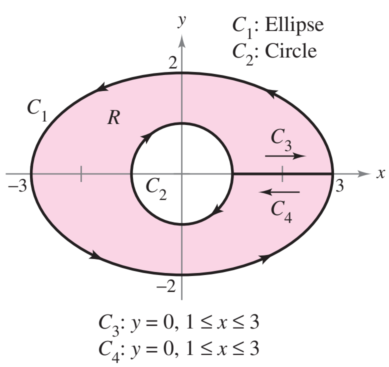


Figure 15.33

Let R be the region inside the ellipse $(x^2/9) + (y^2/4) = 1$ and outside the circle $x^2 + y^2 = 1$. Evaluate the line integral

$$\int_C 2xy\,dx + (x^2 + 2x)\,dy$$

where $C = C_1 + C_2$ is the boundary of R, as shown in Figure 15.33.

Solution To begin, you can introduce the line segments C_3 and C_4, as shown in Figure 15.33. Note that because the curves C_3 and C_4 have opposite orientations, the line integrals over them cancel. Furthermore, you can apply Green's Theorem to the region R using the boundary $C_1 + C_4 + C_2 + C_3$ to obtain

$$\begin{aligned} \int_C 2xy\,dx + (x^2 + 2x)\,dy &= \int_R\!\!\int \left(\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y}\right) dA \\ &= \int_R\!\!\int (2x + 2 - 2x)\,dA \\ &= 2\int_R\!\!\int dA \\ &= 2(\text{area of } R) \\ &= 2(\pi ab - \pi r^2) \\ &= 2[\pi(3)(2) - \pi(1^2)] \\ &= 10\pi. \end{aligned}$$

Try It **Exploration A**

In Section 15.1, a necessary and sufficient condition for conservative vector fields was listed. There, only one direction of the proof was shown. You can now outline the other direction, using Green's Theorem. Let $\mathbf{F}(x, y) = M\mathbf{i} + N\mathbf{j}$ be defined on an open disk R. You want to show that if M and N have continuous first partial derivatives and

$$\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x}$$

then $\mathbf{F}$ is conservative. Suppose that C is a closed path forming the boundary of a connected region lying in R. Then, using the fact that $\partial M/\partial y = \partial N/\partial x$, you can apply Green's Theorem to conclude that

$$\begin{aligned}\int_C \mathbf{F} \cdot d\mathbf{r} &= \int_C M\,dx + N\,dy \\ &= \int\!\!\int_R \left(\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y}\right) dA \\ &= 0.\end{aligned}$$

This, in turn, is equivalent to showing that $\mathbf{F}$ is conservative (see Theorem 15.7).

Alternative Forms of Green's Theorem

This section concludes with the derivation of two vector forms of Green's Theorem for regions in the plane. The extension of these vector forms to three dimensions is the basis for the discussion in the remaining sections of this chapter. If $\mathbf{F}$ is a vector field in the plane, you can write

$$\mathbf{F}(x, y, z) = M\mathbf{i} + N\mathbf{j} + 0\mathbf{k}$$

so that the curl of $\mathbf{F}$, as described in Section 15.1, is given by

$$\begin{aligned}\mathbf{curl\ F} = \nabla \times \mathbf{F} &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ M & N & 0 \end{vmatrix} \\ &= -\frac{\partial N}{\partial z}\mathbf{i} + \frac{\partial M}{\partial z}\mathbf{j} + \left(\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y}\right)\mathbf{k}.\end{aligned}$$

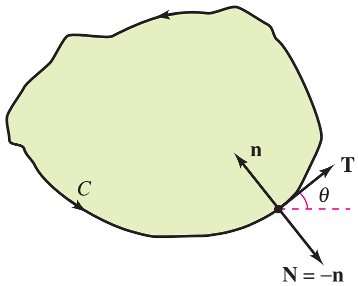
Consequently,

$$\begin{aligned}(\mathbf{curl\ F}) \cdot \mathbf{k} &= \left[-\frac{\partial N}{\partial z}\mathbf{i} + \frac{\partial M}{\partial z}\mathbf{j} + \left(\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y}\right)\mathbf{k}\right] \cdot \mathbf{k} \\ &= \frac{\partial N}{\partial x} - \frac{\partial M}{\partial y}.\end{aligned}$$

With appropriate conditions on $\mathbf{F}$, C, and R, you can write Green's Theorem in the vector form

$$\begin{aligned}\int_C \mathbf{F} \cdot d\mathbf{r} &= \int\!\!\int_R \left(\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y}\right) dA \\ &= \int\!\!\int_R (\mathbf{curl\ F}) \cdot \mathbf{k}\, dA. \qquad \text{First alternative form}\end{aligned}$$

The extension of this vector form of Green's Theorem to surfaces in space produces **Stokes's Theorem,** discussed in Section 15.8.



$\mathbf{T} = \cos\theta\mathbf{i} + \sin\theta\mathbf{j}$

$\mathbf{n} = \cos\left(\theta + \frac{\pi}{2}\right)\mathbf{i} + \sin\left(\theta + \frac{\pi}{2}\right)\mathbf{j}$

$\quad = -\sin\theta\mathbf{i} + \cos\theta\mathbf{j}$

$\mathbf{N} = \sin\theta\mathbf{i} - \cos\theta\mathbf{j}$

Figure 15.34

For the second vector form of Green's Theorem, assume the same conditions for $\mathbf{F}$, C, and R. Using the arc length parameter s for C, you have $\mathbf{r}(s) = x(s)\mathbf{i} + y(s)\mathbf{j}$. So, a unit tangent vector $\mathbf{T}$ to curve C is given by

$$\mathbf{r}'(s) = \mathbf{T} = x'(s)\mathbf{i} + y'(s)\mathbf{j}.$$

From Figure 15.34 you can see that the *outward* unit normal vector $\mathbf{N}$ can then be written as

$$\mathbf{N} = y'(s)\mathbf{i} - x'(s)\mathbf{j}.$$

Consequently, for $\mathbf{F}(x, y) = M\mathbf{i} + N\mathbf{j}$, you can apply Green's Theorem to obtain

$$\begin{aligned}
\int_C \mathbf{F} \cdot \mathbf{N}\, ds &= \int_a^b (M\mathbf{i} + N\mathbf{j}) \cdot (y'(s)\mathbf{i} - x'(s)\mathbf{j})\, ds \\
&= \int_a^b \left(M\frac{dy}{ds} - N\frac{dx}{ds}\right) ds \\
&= \int_C M\, dy - N\, dx \\
&= \int_C -N\, dx + M\, dy \\
&= \int_R\!\!\int \left(\frac{\partial M}{\partial x} + \frac{\partial N}{\partial y}\right) dA && \text{Green's Theorem} \\
&= \int_R\!\!\int \operatorname{div} \mathbf{F}\, dA.
\end{aligned}$$

Therefore,

$$\int_C \mathbf{F} \cdot \mathbf{N}\, ds = \int_R\!\!\int \operatorname{div} \mathbf{F}\, dA. \qquad \text{Second alternative form}$$

The extension of this form to three dimensions is called the **Divergence Theorem,** discussed in Section 15.7. The physical interpretations of divergence and curl will be discussed in Sections 15.7 and 15.8.

Section 15.5 Parametric Surfaces

- **Understand the definition of and sketch a parametric surface.**
- **Find a set of parametric equations to represent a surface.**
- **Find a normal vector and a tangent plane to a parametric surface.**
- **Find the area of a parametric surface.**

Parametric Surfaces

You already know how to represent a curve in the plane or in space by a set of parametric equations—or, equivalently, by a vector-valued function.

$$\mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j} \qquad \text{Plane curve}$$

$$\mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j} + z(t)\mathbf{k} \qquad \text{Space curve}$$

In this section, you will learn how to represent a surface in space by a set of parametric equations—or by a vector-valued function. For curves, note that the vector-valued function $\mathbf{r}$ is a function of a *single* parameter t. For surfaces, the vector-valued function is a function of *two* parameters u and v.

Definition of Parametric Surface

Let x, y, and z be functions of u and v that are continuous on a domain D in the uv-plane. The set of points (x, y, z) given by

$$\mathbf{r}(u, v) = x(u, v)\mathbf{i} + y(u, v)\mathbf{j} + z(u, v)\mathbf{k} \qquad \text{Parametric surface}$$

is called a **parametric surface.** The equations

$$x = x(u, v), \quad y = y(u, v), \quad \text{and} \quad z = z(u, v) \qquad \text{Parametric equations}$$

are the **parametric equations** for the surface.

If S is a parametric surface given by the vector-valued function $\mathbf{r}$, then S is traced out by the position vector $\mathbf{r}(u, v)$ as the point (u, v) moves throughout the domain D, as shown in Figure 15.35.

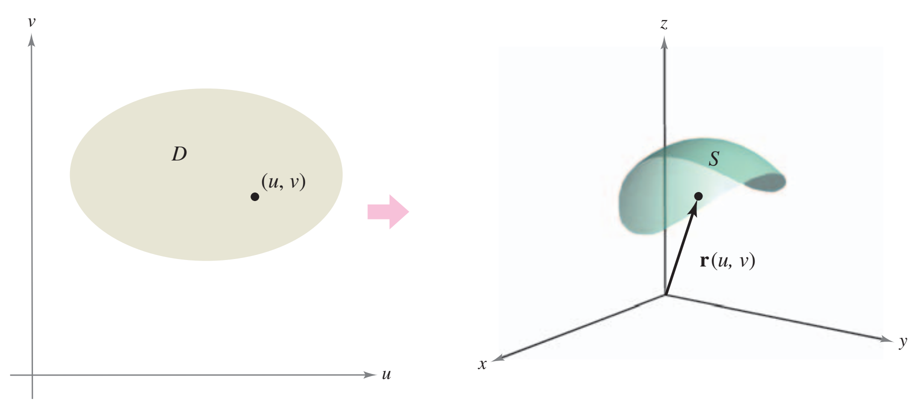


Figure 15.35

Rotatable Graph

TECHNOLOGY Some computer algebra systems are capable of graphing surfaces that are represented parametrically. If you have access to such software, use it to graph some of the surfaces in the examples and exercises in this section.

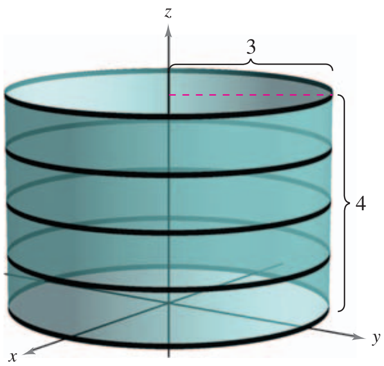


Figure 15.36

Rotatable Graph

EXAMPLE 1 Sketching a Parametric Surface

Identify and sketch the parametric surface S given by

$$\mathbf{r}(u, v) = 3\cos u\mathbf{i} + 3\sin u\mathbf{j} + v\mathbf{k}$$

where $0 \le u \le 2\pi$ and $0 \le v \le 4$.

Solution Because $x = 3\cos u$ and $y = 3\sin u$, you know that for each point (x, y, z) on the surface, x and y are related by the equation $x^2 + y^2 = 3^2$. In other words, each cross section of S taken parallel to the xy-plane is a circle of radius 3, centered on the z-axis. Because $z = v$, where $0 \le v \le 4$, you can see that the surface is a right circular cylinder of height 4. The radius of the cylinder is 3, and the z-axis forms the axis of the cylinder, as shown in Figure 15.36.

Try It Exploration A

As with parametric representations of curves, parametric representations of surfaces are not unique. That is, there are many other sets of parametric equations that could be used to represent the surface shown in Figure 15.36.

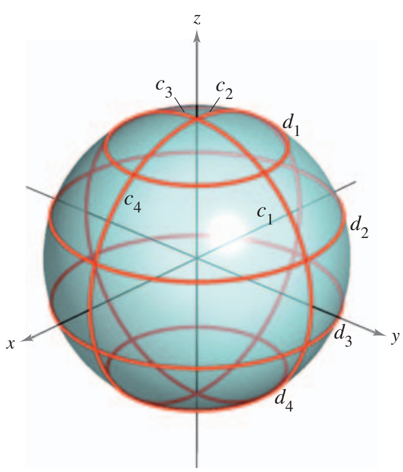


Figure 15.37

Rotatable Graph

EXAMPLE 2 Sketching a Parametric Surface

Identify and sketch the parametric surface S given by

$$\mathbf{r}(u, v) = \sin u \cos v\mathbf{i} + \sin u \sin v\mathbf{j} + \cos u\mathbf{k}$$

where $0 \le u \le \pi$ and $0 \le v \le 2\pi$.

Solution To identify the surface, you can try to use trigonometric identities to eliminate the parameters. After some experimentation, you can discover that

$$\begin{aligned}
x^2 + y^2 + z^2 &= (\sin u \cos v)^2 + (\sin u \sin v)^2 + (\cos u)^2 \\
&= \sin^2 u \cos^2 v + \sin^2 u \sin^2 v + \cos^2 u \\
&= \sin^2 u(\cos^2 v + \sin^2 v) + \cos^2 u \\
&= \sin^2 u + \cos^2 u \\
&= 1.
\end{aligned}$$

So, each point on S lies on the unit sphere, centered at the origin, as shown in Figure 15.37. For fixed $u = d_i$, $\mathbf{r}(u, v)$ traces out latitude circles

$$x^2 + y^2 = \sin^2 d_i, \quad 0 \le d_i \le \pi$$

that are parallel to the xy-plane, and for fixed $v = c_i$, $\mathbf{r}(u, v)$ traces out longitude (or meridian) half-circles.

Try It Exploration A

NOTE To convince yourself further that the vector-valued function in Example 2 traces out the entire unit sphere, recall that the parametric equations

$$x = \rho \sin\phi \cos\theta, \quad y = \rho \sin\phi \sin\theta, \quad \text{and} \quad z = \rho\cos\phi$$

where $0 \le \theta \le 2\pi$ and $0 \le \phi \le \pi$, describe the conversion from spherical to rectangular coordinates, as discussed in Section 11.7.

Finding Parametric Equations for Surfaces

In Examples 1 and 2, you were asked to identify the surface described by a given set of parametric equations. The reverse problem—that of writing a set of parametric equations for a given surface—is generally more difficult. One type of surface for which this problem is straightforward, however, is a surface that is given by $z = f(x, y)$. You can parametrize such a surface as

$$\mathbf{r}(x, y) = x\mathbf{i} + y\mathbf{j} + f(x, y)\mathbf{k}.$$

EXAMPLE 3 Representing a Surface Parametrically

Write a set of parametric equations for the cone given by

$$z = \sqrt{x^2 + y^2}$$

as shown in Figure 15.38.

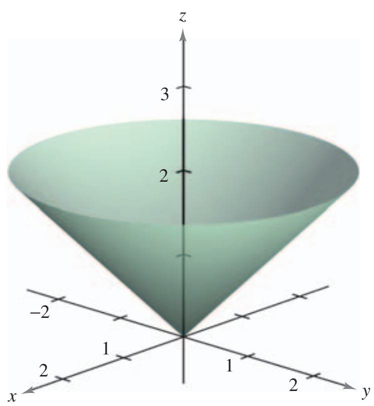


Figure 15.38

Rotatable Graph

Solution Because this surface is given in the form $z = f(x, y)$, you can let x and y be the parameters. Then the cone is represented by the vector-valued function

$$\mathbf{r}(x, y) = x\mathbf{i} + y\mathbf{j} + \sqrt{x^2 + y^2}\,\mathbf{k}$$

where (x, y) varies over the entire xy-plane.

Try It Exploration A

A second type of surface that is easily represented parametrically is a surface of revolution. For instance, to represent the surface formed by revolving the graph of $y = f(x)$, $a \le x \le b$, about the x-axis, use

$$x = u, \quad y = f(u)\cos v, \quad \text{and} \quad z = f(u)\sin v$$

where $a \le u \le b$ and $0 \le v \le 2\pi$. Select the animation button to see an example.

Animation

EXAMPLE 4 Representing a Surface of Revolution Parametrically

Write a set of parametric equations for the surface of revolution obtained by revolving

$$f(x) = \frac{1}{x}, \quad 1 \le x \le 10$$

about the x-axis.

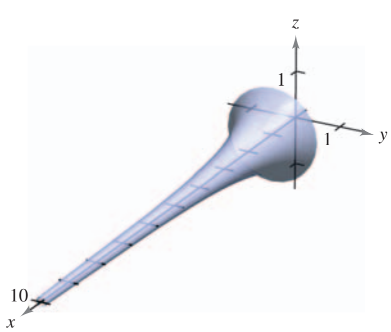


Figure 15.39

Rotatable Graph

Solution Use the parameters u and v as described above to write

$$x = u, \quad y = f(u)\cos v = \frac{1}{u}\cos v, \quad \text{and} \quad z = f(u)\sin v = \frac{1}{u}\sin v$$

where $1 \le u \le 10$ and $0 \le v \le 2\pi$. The resulting surface is a portion of *Gabriel's Horn*, as shown in Figure 15.39.

Try It Exploration A Open Exploration

The surface of revolution in Example 4 is formed by revolving the graph of $y = f(x)$ about the x-axis. For other types of surfaces of revolution, a similar parametrization can be used. For instance, to parametrize the surface formed by revolving the graph of $x = f(z)$ about the z-axis, you can use

$$z = u, \quad x = f(u)\cos v, \quad \text{and} \quad y = f(u)\sin v.$$

Normal Vectors and Tangent Planes

Let S be a parametric surface given by

$$\mathbf{r}(u, v) = x(u, v)\mathbf{i} + y(u, v)\mathbf{j} + z(u, v)\mathbf{k}$$

over an open region D such that x, y, and z have continuous partial derivatives on D. The **partial derivatives of r** with respect to u and v are defined as

$$\mathbf{r}_u = \frac{\partial x}{\partial u}(u, v)\mathbf{i} + \frac{\partial y}{\partial u}(u, v)\mathbf{j} + \frac{\partial z}{\partial u}(u, v)\mathbf{k}$$

and

$$\mathbf{r}_v = \frac{\partial x}{\partial v}(u, v)\mathbf{i} + \frac{\partial y}{\partial v}(u, v)\mathbf{j} + \frac{\partial z}{\partial v}(u, v)\mathbf{k}.$$

Each of these partial derivatives is a vector-valued function that can be interpreted geometrically in terms of tangent vectors. For instance, if $v = v_0$ is held constant, then $\mathbf{r}(u, v_0)$ is a vector-valued function of a single parameter and defines a curve C_1 that lies on the surface S. The tangent vector to C_1 at the point $(x(u_0, v_0),\ y(u_0, v_0),\ z(u_0, v_0))$ is given by

$$\mathbf{r}_u(u_0, v_0) = \frac{\partial x}{\partial u}(u_0, v_0)\mathbf{i} + \frac{\partial y}{\partial u}(u_0, v_0)\mathbf{j} + \frac{\partial z}{\partial u}(u_0, v_0)\mathbf{k}$$

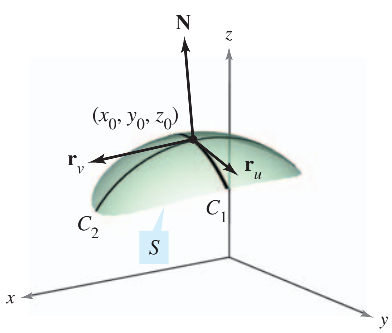


Figure 15.40

Rotatable Graph

as shown in Figure 15.40. In a similar way, if $u = u_0$ is held constant, then $\mathbf{r}(u_0, v)$ is a vector-valued function of a single parameter and defines a curve C_2 that lies on the surface S. The tangent vector to C_2 at the point $(x(u_0, v_0), y(u_0, v_0), z(u_0, v_0))$ is given by

$$\mathbf{r}_v(u_0, v_0) = \frac{\partial x}{\partial v}(u_0, v_0)\mathbf{i} + \frac{\partial y}{\partial v}(u_0, v_0)\mathbf{j} + \frac{\partial z}{\partial v}(u_0, v_0)\mathbf{k}.$$

If the normal vector $\mathbf{r}_u \times \mathbf{r}_v$ is not $\mathbf{0}$ for any (u, v) in D, the surface S is called **smooth** and will have a tangent plane. Informally, a smooth surface is one that has no sharp points or cusps. For instance, spheres, ellipsoids, and paraboloids are smooth, whereas the cone given in Example 3 is not smooth.

Normal Vector to a Smooth Parametric Surface

Let S be a smooth parametric surface

$$\mathbf{r}(u, v) = x(u, v)\mathbf{i} + y(u, v)\mathbf{j} + z(u, v)\mathbf{k}$$

defined over an open region D in the uv-plane. Let (u_0, v_0) be a point in D. A normal vector at the point

$$(x_0, y_0, z_0) = (x(u_0, v_0), y(u_0, v_0), z(u_0, v_0))$$

is given by

$$\mathbf{N} = \mathbf{r}_u(u_0, v_0) \times \mathbf{r}_v(u_0, v_0) = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \dfrac{\partial x}{\partial u} & \dfrac{\partial y}{\partial u} & \dfrac{\partial z}{\partial u} \\ \dfrac{\partial x}{\partial v} & \dfrac{\partial y}{\partial v} & \dfrac{\partial z}{\partial v} \end{vmatrix}.$$

NOTE Figure 15.40 shows the normal vector $\mathbf{r}_u \times \mathbf{r}_v$. The vector $\mathbf{r}_v \times \mathbf{r}_u$ is also normal to S and points in the opposite direction.

EXAMPLE 5 Finding a Tangent Plane to a Parametric Surface

Find an equation of the tangent plane to the paraboloid given by

$$\mathbf{r}(u, v) = u\mathbf{i} + v\mathbf{j} + (u^2 + v^2)\mathbf{k}$$

at the point $(1, 2, 5)$.

Solution The point in the uv-plane that is mapped to the point $(x, y, z) = (1, 2, 5)$ is $(u, v) = (1, 2)$. The partial derivatives of $\mathbf{r}$ are

$$\mathbf{r}_u = \mathbf{i} + 2u\mathbf{k} \quad \text{and} \quad \mathbf{r}_v = \mathbf{j} + 2v\mathbf{k}.$$

The normal vector is given by

$$\mathbf{r}_u \times \mathbf{r}_v = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & 0 & 2u \\ 0 & 1 & 2v \end{vmatrix} = -2u\mathbf{i} - 2v\mathbf{j} + \mathbf{k}$$

which implies that the normal vector at $(1, 2, 5)$ is $\mathbf{r}_u \times \mathbf{r}_v = -2\mathbf{i} - 4\mathbf{j} + \mathbf{k}$. So, an equation of the tangent plane at $(1, 2, 5)$ is

$$-2(x - 1) - 4(y - 2) + (z - 5) = 0$$
$$-2x - 4y + z = -5.$$

The tangent plane is shown in Figure 15.41.

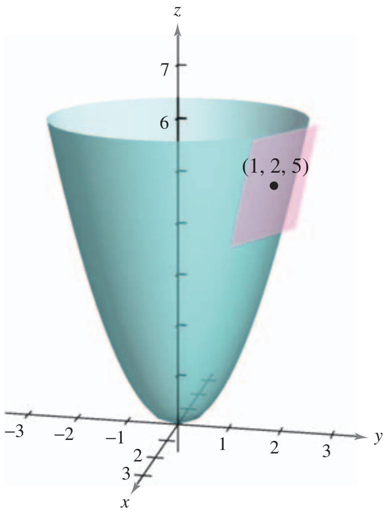


Figure 15.41

Rotatable Graph

Try It

Exploration A

Area of a Parametric Surface

To define the area of a parametric surface, you can use a development that is similar to that given in Section 14.5. Begin by constructing an inner partition of D consisting of n rectangles, where the area of the ith rectangle D_i is $\Delta A_i = \Delta u_i \, \Delta v_i$, as shown in Figure 15.42. In each D_i let (u_i, v_i) be the point that is closest to the origin. At the point $(x_i, y_i, z_i) = (x(u_i, v_i), y(u_i, v_i), z(u_i, v_i))$ on the surface S, construct a tangent plane T_i. The area of the portion of S that corresponds to D_i, ΔT_i, can be approximated by a parallelogram in the tangent plane. That is, $\Delta T_i \approx \Delta S_i$. So, the surface of S is given by $\Sigma \, \Delta S_i \approx \Sigma \, \Delta T_i$. The area of the parallelogram in the tangent plane is

$$\|\Delta u_i \mathbf{r}_u \times \Delta v_i \mathbf{r}_v\| = \|\mathbf{r}_u \times \mathbf{r}_v\| \, \Delta u_i \, \Delta v_i$$

which leads to the following definition.

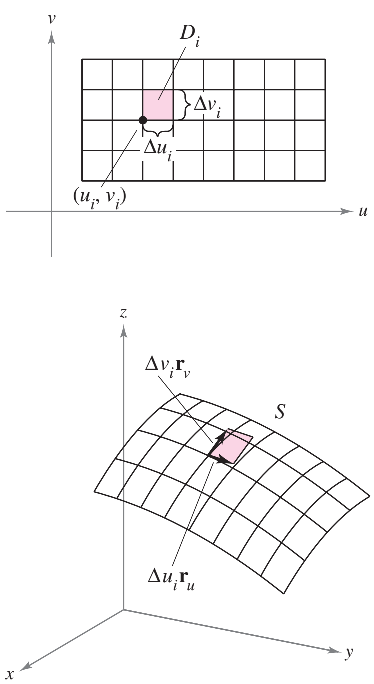


Figure 15.42

Area of a Parametric Surface

Let S be a smooth parametric surface

$$\mathbf{r}(u, v) = x(u, v)\mathbf{i} + y(u, v)\mathbf{j} + z(u, v)\mathbf{k}$$

defined over an open region D in the uv-plane. If each point on the surface S corresponds to exactly one point in the domain D, then the **surface area** of S is given by

$$\text{Surface area} = \iint_S dS = \iint_D \|\mathbf{r}_u \times \mathbf{r}_v\| \, dA$$

where $\mathbf{r}_u = \dfrac{\partial x}{\partial u}\mathbf{i} + \dfrac{\partial y}{\partial u}\mathbf{j} + \dfrac{\partial z}{\partial u}\mathbf{k}$ and $\mathbf{r}_v = \dfrac{\partial x}{\partial v}\mathbf{i} + \dfrac{\partial y}{\partial v}\mathbf{j} + \dfrac{\partial z}{\partial v}\mathbf{k}$.

For a surface S given by $z = f(x, y)$, this formula for surface area corresponds to that given in Section 14.5. To see this, you can parametrize the surface using the vector-valued function

$$\mathbf{r}(x, y) = x\mathbf{i} + y\mathbf{j} + f(x, y)\mathbf{k}$$

defined over the region R in the xy-plane. Using

$$\mathbf{r}_x = \mathbf{i} + f_x(x, y)\mathbf{k} \quad \text{and} \quad \mathbf{r}_y = \mathbf{j} + f_y(x, y)\mathbf{k}$$

you have

$$\mathbf{r}_x \times \mathbf{r}_y = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & 0 & f_x(x, y) \\ 0 & 1 & f_y(x, y) \end{vmatrix} = -f_x(x, y)\mathbf{i} - f_y(x, y)\mathbf{j} + \mathbf{k}$$

and $\|\mathbf{r}_x \times \mathbf{r}_y\| = \sqrt{[f_x(x, y)]^2 + [f_y(x, y)]^2 + 1}$. This implies that the surface area of S is

$$\begin{aligned} \text{Surface area} &= \iint_R \|\mathbf{r}_x \times \mathbf{r}_y\|\, dA \\ &= \iint_R \sqrt{1 + [f_x(x, y)]^2 + [f_y(x, y)]^2}\, dA. \end{aligned}$$

EXAMPLE 6 Finding Surface Area

Find the surface area of the unit sphere given by

$$\mathbf{r}(u, v) = \sin u \cos v\mathbf{i} + \sin u \sin v\mathbf{j} + \cos u\mathbf{k}$$

where the domain D is given by $0 \le u \le \pi$ and $0 \le v \le 2\pi$.

NOTE The surface in Example 6 does not quite fulfill the hypothesis that each point on the surface corresponds to exactly one point in D. For this surface, $\mathbf{r}(u, 0) = \mathbf{r}(u, 2\pi)$ for any fixed value of u. However, because the overlap consists of only a semicircle (which has no area), you can still apply the formula for the area of a parametric surface.

Solution Begin by calculating $\mathbf{r}_u$ and $\mathbf{r}_v$.

$$\begin{aligned} \mathbf{r}_u &= \cos u \cos v\mathbf{i} + \cos u \sin v\mathbf{j} - \sin u\mathbf{k} \\ \mathbf{r}_v &= -\sin u \sin v\mathbf{i} + \sin u \cos v\mathbf{j} \end{aligned}$$

The cross product of these two vectors is

$$\begin{aligned} \mathbf{r}_u \times \mathbf{r}_v &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \cos u \cos v & \cos u \sin v & -\sin u \\ -\sin u \sin v & \sin u \cos v & 0 \end{vmatrix} \\ &= \sin^2 u \cos v\mathbf{i} + \sin^2 u \sin v\mathbf{j} + \sin u \cos u\mathbf{k} \end{aligned}$$

which implies that

$$\begin{aligned} \|\mathbf{r}_u \times \mathbf{r}_v\| &= \sqrt{(\sin^2 u \cos v)^2 + (\sin^2 u \sin v)^2 + (\sin u \cos u)^2} \\ &= \sqrt{\sin^4 u + \sin^2 u \cos^2 u} \\ &= \sqrt{\sin^2 u} \\ &= \sin u. \qquad \sin u > 0 \text{ for } 0 \le u \le \pi \end{aligned}$$

Finally, the surface area of the sphere is

$$\begin{aligned} A = \iint_D \|\mathbf{r}_u \times \mathbf{r}_v\|\, dA &= \int_0^{2\pi}\int_0^{\pi} \sin u\, du\, dv \\ &= \int_0^{2\pi} 2\, dv \\ &= 4\pi. \end{aligned}$$

Try It **Exploration A**

EXPLORATION

For the torus in Example 7, describe the function $\mathbf{r}(u, v)$ for fixed u. Then describe the function $\mathbf{r}(u, v)$ for fixed v.

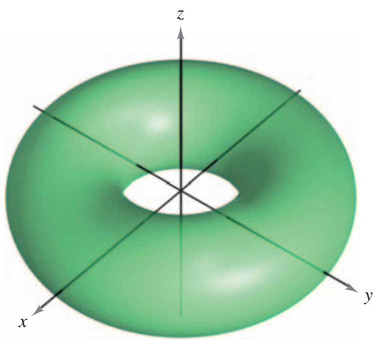


Figure 15.43

Rotatable Graph

EXAMPLE 7 Finding Surface Area

Find the surface area of the torus given by

$$\mathbf{r}(u, v) = (2 + \cos u)\cos v\mathbf{i} + (2 + \cos u)\sin v\mathbf{j} + \sin u\mathbf{k}$$

where the domain D is given by $0 \le u \le 2\pi$ and $0 \le v \le 2\pi$. (See Figure 15.43.)

Solution Begin by calculating $\mathbf{r}_u$ and $\mathbf{r}_v$.

$$\mathbf{r}_u = -\sin u\cos v\mathbf{i} - \sin u\sin v\mathbf{j} + \cos u\mathbf{k}$$
$$\mathbf{r}_v = -(2 + \cos u)\sin v\mathbf{i} + (2 + \cos u)\cos v\mathbf{j}$$

The cross product of these two vectors is

$$\begin{aligned}
\mathbf{r}_u \times \mathbf{r}_v &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ -\sin u\cos v & -\sin u\sin v & \cos u \\ -(2 + \cos u)\sin v & (2 + \cos u)\cos v & 0 \end{vmatrix} \\
&= -(2 + \cos u)(\cos v\cos u\mathbf{i} + \sin v\cos u\mathbf{j} + \sin u\mathbf{k})
\end{aligned}$$

which implies that

$$\begin{aligned}
\|\mathbf{r}_u \times \mathbf{r}_v\| &= (2 + \cos u)\sqrt{(\cos v\cos u)^2 + (\sin v\cos u)^2 + \sin^2 u} \\
&= (2 + \cos u)\sqrt{\cos^2 u(\cos^2 v + \sin^2 v) + \sin^2 u} \\
&= (2 + \cos u)\sqrt{\cos^2 u + \sin^2 u} \\
&= 2 + \cos u.
\end{aligned}$$

Finally, the surface area of the torus is

$$\begin{aligned}
A = \iint_D \|\mathbf{r}_u \times \mathbf{r}_v\|\, dA &= \int_0^{2\pi}\int_0^{2\pi} (2 + \cos u)\, du\, dv \\
&= \int_0^{2\pi} 4\pi\, dv \\
&= 8\pi^2.
\end{aligned}$$

Try It | Exploration A | Exploration B

If the surface S is a surface of revolution, you can show that the formula for surface area given in Section 7.4 is equivalent to the formula given in this section. For instance, suppose f is a nonnegative function such that f' is continuous over the interval $[a, b]$. Let S be the surface of revolution formed by revolving the graph of f, where $a \le x \le b$, about the x-axis. From Section 7.4, you know that the surface area is given by

$$\text{Surface area} = 2\pi\int_a^b f(x)\sqrt{1 + [f'(x)]^2}\, dx.$$

To represent S parametrically, let $x = u$, $y = f(u)\cos v$, and $z = f(u)\sin v$, where $a \le u \le b$ and $0 \le v \le 2\pi$. Then,

$$\mathbf{r}(u, v) = u\mathbf{i} + f(u)\cos v\mathbf{j} + f(u)\sin v\mathbf{k}.$$

Try showing that the formula

$$\text{Surface area} = \iint_D \|\mathbf{r}_u \times \mathbf{r}_v\|\, dA$$

is equivalent to the formula given above (see Exercise 52).

Section 15.6 Surface Integrals

- **Evaluate a surface integral as a double integral.**
- **Evaluate a surface integral for a parametric surface.**
- **Determine the orientation of a surface.**
- **Understand the concept of a flux integral.**

Surface Integrals

The remainder of this chapter deals primarily with **surface integrals.** You will first consider surfaces given by $z = g(x, y)$. Later in this section you will consider more general surfaces given in parametric form.

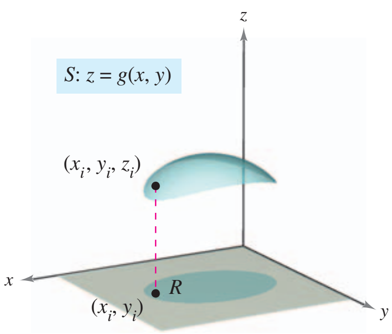
Let S be a surface given by $z = g(x, y)$ and let R be its projection onto the xy-plane, as shown in Figure 15.44. Suppose that g, g_x, and g_y are continuous at all points in R and that f is defined on S. Employing the procedure used to find surface area in Section 14.5, evaluate f at (x_i, y_i, z_i) and form the sum

$$\sum_{i=1}^{n} f(x_i, y_i, z_i)\,\Delta S_i$$

where $\Delta S_i \approx \sqrt{1 + [g_x(x_i, y_i)]^2 + [g_y(x_i, y_i)]^2}\,\Delta A_i$. Provided the limit of the above sum as $\|\Delta\|$ approaches 0 exists, the **surface integral of f over S** is defined as

$$\int_S\!\!\int f(x, y, z)\,dS = \lim_{\|\Delta\|\to 0} \sum_{i=1}^{n} f(x_i, y_i, z_i)\,\Delta S_i.$$

This integral can be evaluated by a double integral.



Scalar function f assigns a number to each point of S.
Figure 15.44

Rotatable Graph

THEOREM 15.10 Evaluating a Surface Integral

Let S be a surface with equation $z = g(x, y)$ and let R be its projection onto the xy-plane. If g, g_x, and g_y are continuous on R and f is continuous on S, then the surface integral of f over S is

$$\int_S\!\!\int f(x, y, z)\,dS = \int_R\!\!\int f(x, y, g(x, y))\sqrt{1 + [g_x(x, y)]^2 + [g_y(x, y)]^2}\,dA.$$

For surfaces described by functions of x and z (or y and z), you can make the following adjustments to Theorem 15.10. If S is the graph of $y = g(x, z)$ and R is its projection onto the xz-plane, then

$$\int_S\!\!\int f(x, y, z)\,dS = \int_R\!\!\int f(x, g(x, z), z)\sqrt{1 + [g_x(x, z)]^2 + [g_z(x, z)]^2}\,dA.$$

If S is the graph of $x = g(y, z)$ and R is its projection onto the yz-plane, then

$$\int_S\!\!\int f(x, y, z)\,dS = \int_R\!\!\int f(g(y, z), y, z)\sqrt{1 + [g_y(y, z)]^2 + [g_z(y, z)]^2}\,dA.$$

If $f(x, y, z) = 1$, the surface integral over S yields the surface area of S. For instance, suppose the surface S is the plane given by $z = x$, where $0 \le x \le 1$ and $0 \le y \le 1$. The surface area of S is $\sqrt{2}$ square units. Try verifying that $\int_S\!\int f(x, y, z)\,dS = \sqrt{2}$.

EXAMPLE 1 Evaluating a Surface Integral

Evaluate the surface integral

$$\int_S\!\!\int (y^2 + 2yz)\, dS$$

where S is the first-octant portion of the plane $2x + y + 2z = 6$.

Solution Begin by writing S as

$$z = \frac{1}{2}(6 - 2x - y)$$

$$g(x, y) = \frac{1}{2}(6 - 2x - y).$$

Using the partial derivatives $g_x(x, y) = -1$ and $g_y(x, y) = -\frac{1}{2}$, you can write

$$\sqrt{1 + [g_x(x, y)]^2 + [g_y(x, y)]^2} = \sqrt{1 + 1 + \frac{1}{4}} = \frac{3}{2}.$$

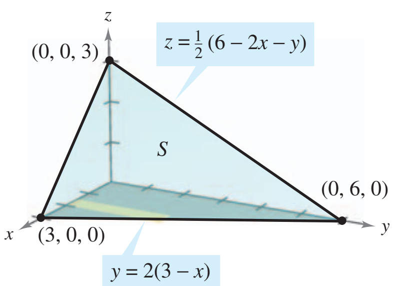


Figure 15.45

Rotatable Graph

Using Figure 15.45 and Theorem 15.10, you obtain

$$\begin{aligned}
\int_S\!\!\int (y^2 + 2yz)\, dS &= \int_R\!\!\int f(x, y, g(x, y))\sqrt{1 + [g_x(x, y)]^2 + [g_y(x, y)]^2}\, dA \\
&= \int_R\!\!\int \left[y^2 + 2y\left(\frac{1}{2}\right)(6 - 2x - y)\right]\left(\frac{3}{2}\right) dA \\
&= 3\int_0^3\!\int_0^{2(3-x)} y(3 - x)\, dy\, dx \\
&= 6\int_0^3 (3 - x)^3\, dx \\
&= -\frac{3}{2}(3 - x)^4\Big]_0^3 \\
&= \frac{243}{2}.
\end{aligned}$$

Try It | Exploration A | Exploration B

An alternative solution to Example 1 would be to project S onto the yz-plane, as shown in Figure 15.46. Then, $x = \frac{1}{2}(6 - y - 2z)$, and

$$\sqrt{1 + [g_y(y, z)]^2 + [g_z(y, z)]^2} = \sqrt{1 + \frac{1}{4} + 1} = \frac{3}{2}.$$

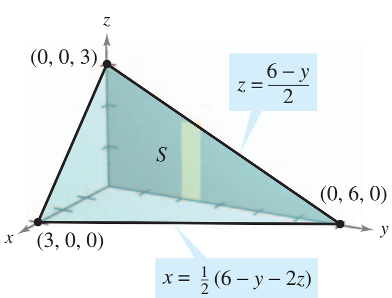


Figure 15.46

Rotatable Graph

So, the surface integral is

$$\begin{aligned}
\int_S\!\!\int (y^2 + 2yz)\, dS &= \int_R\!\!\int f(g(y, z), y, z)\sqrt{1 + [g_y(y, z)]^2 + [g_z(y, z)]^2}\, dA \\
&= \int_0^6\!\int_0^{(6-y)/2} (y^2 + 2yz)\left(\frac{3}{2}\right) dz\, dy \\
&= \frac{3}{8}\int_0^6 (36y - y^3)\, dy \\
&= \frac{243}{2}.
\end{aligned}$$

Try reworking Example 1 by projecting S onto the xz-plane.

In Example 1, you could have projected the surface S onto any one of the three coordinate planes. In Example 2, S is a portion of a cylinder centered about the x-axis, and you can project it onto either the xz-plane or the xy-plane.

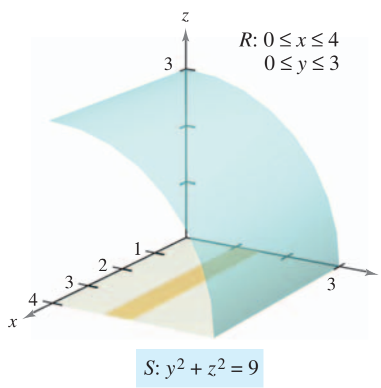


Figure 15.47

Rotatable Graph

EXAMPLE 2 Evaluating a Surface Integral

Evaluate the surface integral

$$\iint_S (x + z)\, dS$$

where S is the first-octant portion of the cylinder $y^2 + z^2 = 9$ between $x = 0$ and $x = 4$, as shown in Figure 15.47.

Solution Project S onto the xy-plane, so that $z = g(x, y) = \sqrt{9 - y^2}$, and obtain

$$\sqrt{1 + [g_x(x, y)]^2 + [g_y(x, y)]^2} = \sqrt{1 + \left(\frac{-y}{\sqrt{9 - y^2}}\right)^2} = \frac{3}{\sqrt{9 - y^2}}.$$

Theorem 15.10 does not apply directly because g_y is not continuous when $y = 3$. However, you can apply the theorem for $0 \le b < 3$ and then take the limit as b approaches 3, as follows.

$$\begin{aligned}
\iint_S (x + z)\, dS &= \lim_{b \to 3^-} \int_0^b \int_0^4 \left(x + \sqrt{9 - y^2}\right) \frac{3}{\sqrt{9 - y^2}}\, dx\, dy \\
&= \lim_{b \to 3^-} 3 \int_0^b \int_0^4 \left(\frac{x}{\sqrt{9 - y^2}} + 1\right) dx\, dy \\
&= \lim_{b \to 3^-} 3 \int_0^b \left[\frac{x^2}{2\sqrt{9 - y^2}} + x\right]_0^4 dy \\
&= \lim_{b \to 3^-} 3 \int_0^b \left(\frac{8}{\sqrt{9 - y^2}} + 4\right) dy \\
&= \lim_{b \to 3^-} 3 \left[4y + 8 \arcsin \frac{y}{3}\right]_0^b \\
&= \lim_{b \to 3^-} 3 \left(4b + 8 \arcsin \frac{b}{3}\right) \\
&= 36 + 24\left(\frac{\pi}{2}\right) \\
&= 36 + 12\pi
\end{aligned}$$

Try It | Exploration A | Exploration B | Open Exploration

TECHNOLOGY Some computer algebra systems are capable of evaluating improper integrals. If you have access to such computer software, use it to evaluate the improper integral

$$\int_0^3 \int_0^4 \left(x + \sqrt{9 - y^2}\right) \frac{3}{\sqrt{9 - y^2}}\, dx\, dy.$$

Do you obtain the same result as in Example 2?

You have already seen that if the function f defined on the surface S is simply $f(x, y, z) = 1$, the surface integral yields the *surface area* of S.

$$\text{Area of surface} = \int\!\!\int_S 1\, dS$$

On the other hand, if S is a lamina of variable density and $\rho(x, y, z)$ is the density at the point (x, y, z), then the *mass* of the lamina is given by

$$\text{Mass of lamina} = \int\!\!\int_S \rho(x, y, z)\, dS.$$

EXAMPLE 3 Finding the Mass of a Surface Lamina

A cone-shaped surface lamina S is given by

$$z = 4 - 2\sqrt{x^2 + y^2}, \quad 0 \le z \le 4$$

as shown in Figure 15.48. At each point on S, the density is proportional to the distance between the point and the z-axis. Find the mass m of the lamina.

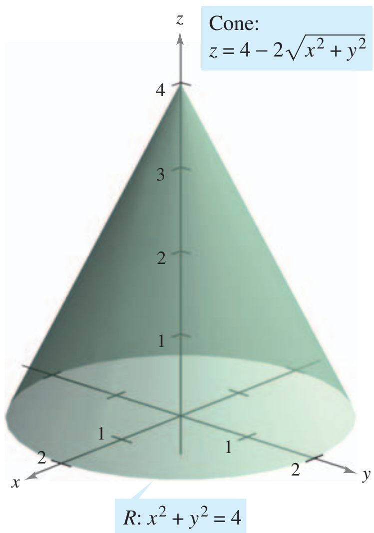


Figure 15.48

Rotatable Graph

Solution Projecting S onto the xy-plane produces

$$S: z = 4 - 2\sqrt{x^2 + y^2} = g(x, y), \quad 0 \le z \le 4$$
$$R: x^2 + y^2 \le 4$$

with a density of $\rho(x, y, z) = k\sqrt{x^2 + y^2}$. Using a surface integral, you can find the mass to be

$$\begin{aligned}
m &= \int\!\!\int_S \rho(x, y, z)\, dS \\
&= \int\!\!\int_R k\sqrt{x^2 + y^2}\sqrt{1 + [g_x(x, y)]^2 + [g_y(x, y)]^2}\, dA \\
&= k\int\!\!\int_R \sqrt{x^2 + y^2}\sqrt{1 + \frac{4x^2}{x^2 + y^2} + \frac{4y^2}{x^2 + y^2}}\, dA \\
&= k\int\!\!\int_R \sqrt{5}\sqrt{x^2 + y^2}\, dA \\
&= k\int_0^{2\pi}\int_0^2 \left(\sqrt{5}r\right)r\, dr\, d\theta && \text{Polar coordinates} \\
&= \frac{\sqrt{5}k}{3}\int_0^{2\pi} r^3\Big]_0^2\, d\theta \\
&= \frac{8\sqrt{5}k}{3}\int_0^{2\pi} d\theta \\
&= \frac{8\sqrt{5}k}{3}\Big[\theta\Big]_0^{2\pi} = \frac{16\sqrt{5}k\pi}{3}.
\end{aligned}$$

Try It Exploration A Exploration B

TECHNOLOGY Use a computer algebra system to confirm the result shown in Example 3. The computer algebra system *Derive* evaluated the integral as follows.

$$k\int_{-2}^{2}\int_{-\sqrt{4-y^2}}^{\sqrt{4-y^2}} \sqrt{5}\sqrt{x^2 + y^2}\, dx\, dy = \frac{16\sqrt{5}k\pi}{3}$$

Parametric Surfaces and Surface Integrals

For a surface S given by the vector-valued function

$$\mathbf{r}(u, v) = x(u, v)\mathbf{i} + y(u, v)\mathbf{j} + z(u, v)\mathbf{k} \qquad \text{Parametric surface}$$

defined over a region D in the uv-plane, you can show that the surface integral of $f(x, y, z)$ over S is given by

$$\int_S\int f(x, y, z)\, dS = \int_D\int f(x(u, v), y(u, v), z(u, v))\, \|\mathbf{r}_u(u, v) \times \mathbf{r}_v(u, v)\|\, dA.$$

Note the similarity to a line integral over a space curve C.

$$\int_C f(x, y, z)\, ds = \int_a^b f(x(t), y(t), z(t))\, \|\mathbf{r}'(t)\|\, dt \qquad \text{Line integral}$$

NOTE Notice that ds and dS can be written as $ds = \|\mathbf{r}'(t)\|\, dt$ and $dS = \|\mathbf{r}_u(u, v) \times \mathbf{r}_v(u, v)\|\, dA$.

EXAMPLE 4 Evaluating a Surface Integral

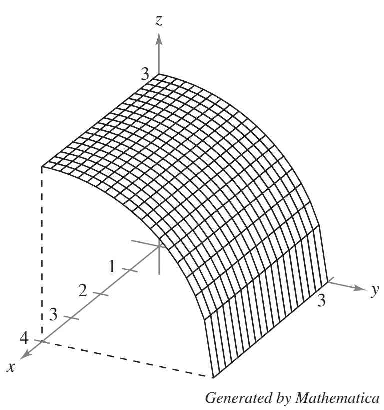


Figure 15.49

Example 2 demonstrated an evaluation of the surface integral

$$\int_S\int (x + z)\, dS$$

where S is the first-octant portion of the cylinder $y^2 + z^2 = 9$ between $x = 0$ and $x = 4$ (see Figure 15.49). Reevaluate this integral in parametric form.

Solution In parametric form, the surface is given by

$$\mathbf{r}(x, \theta) = x\mathbf{i} + 3\cos\theta\mathbf{j} + 3\sin\theta\mathbf{k}$$

where $0 \le x \le 4$ and $0 \le \theta \le \pi/2$. To evaluate the surface integral in parametric form, begin by calculating the following.

$$\mathbf{r}_x = \mathbf{i}$$

$$\mathbf{r}_\theta = -3\sin\theta\mathbf{j} + 3\cos\theta\mathbf{k}$$

$$\mathbf{r}_x \times \mathbf{r}_\theta = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & 0 & 0 \\ 0 & -3\sin\theta & 3\cos\theta \end{vmatrix} = -3\cos\theta\mathbf{j} - 3\sin\theta\mathbf{k}$$

$$\|\mathbf{r}_x \times \mathbf{r}_\theta\| = \sqrt{9\cos^2\theta + 9\sin^2\theta} = 3$$

So, the surface integral can be evaluated as follows.

$$\begin{aligned}\int_D\int (x + 3\sin\theta)3\, dA &= \int_0^4\int_0^{\pi/2} (3x + 9\sin\theta)\, d\theta\, dx \\ &= \int_0^4 \Big[3x\theta - 9\cos\theta\Big]_0^{\pi/2} dx \\ &= \int_0^4 \left(\frac{3\pi}{2}x + 9\right) dx \\ &= \left[\frac{3\pi}{4}x^2 + 9x\right]_0^4 \\ &= 12\pi + 36\end{aligned}$$

Try It

Exploration A

Orientation of a Surface

Unit normal vectors are used to induce an orientation to a surface S in space. A surface is called **orientable** if a unit normal vector $\mathbf{N}$ can be defined at every nonboundary point of S in such a way that the normal vectors vary continuously over the surface S. If this is possible, S is called an **oriented surface.**

An orientable surface S has two distinct sides. So, when you orient a surface, you are selecting one of the two possible unit normal vectors. If S is a closed surface such as a sphere, it is customary to choose the unit normal vector $\mathbf{N}$ to be the one that points outward from the sphere.

Most common surfaces, such as spheres, paraboloids, ellipses, and planes, are orientable. (See Exercise 43 for an example of a surface that is *not* orientable.) Moreover, for an orientable surface, the gradient vector provides a convenient way to find a unit normal vector. That is, for an orientable surface S given by

$$z = g(x, y) \qquad \text{Orientable surface}$$

let

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

Then, S can be oriented by either the unit normal vector

$$\begin{aligned}\mathbf{N} &= \frac{\nabla G(x, y, z)}{\|\nabla G(x, y, z)\|}\\ &= \frac{-g_x(x, y)\mathbf{i} - g_y(x, y)\mathbf{j} + \mathbf{k}}{\sqrt{1 + [g_x(x, y)]^2 + [g_y(x, y)]^2}}\end{aligned} \qquad \text{Upward unit normal}$$

or the unit normal vector

$$\begin{aligned}\mathbf{N} &= \frac{-\nabla G(x, y, z)}{\|\nabla G(x, y, z)\|}\\ &= \frac{g_x(x, y)\mathbf{i} + g_y(x, y)\mathbf{j} - \mathbf{k}}{\sqrt{1 + [g_x(x, y)]^2 + [g_y(x, y)]^2}}\end{aligned} \qquad \text{Downward unit normal}$$

as shown in Figure 15.50. If the smooth orientable surface S is given in parametric form by

$$\mathbf{r}(u, v) = x(u, v)\mathbf{i} + y(u, v)\mathbf{j} + z(u, v)\mathbf{k} \qquad \text{Parametric surface}$$

the unit normal vectors are given by

$$\mathbf{N} = \frac{\mathbf{r}_u \times \mathbf{r}_v}{\|\mathbf{r}_u \times \mathbf{r}_v\|}$$

and

$$\mathbf{N} = \frac{\mathbf{r}_v \times \mathbf{r}_u}{\|\mathbf{r}_v \times \mathbf{r}_u\|}.$$

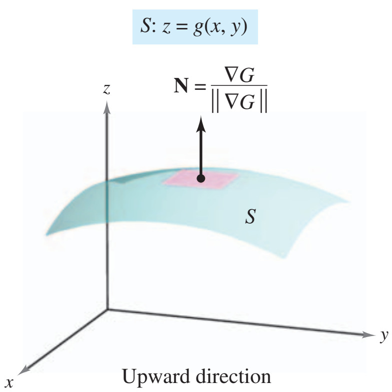
NOTE Suppose that the orientable surface is given by $y = g(x, z)$ or $x = g(y, z)$. Then you can use the gradient vector

$$\nabla G(x, y, z) = -g_x(x, z)\mathbf{i} + \mathbf{j} - g_z(x, z)\mathbf{k} \qquad G(x, y, z) = y - g(x, z)$$

or

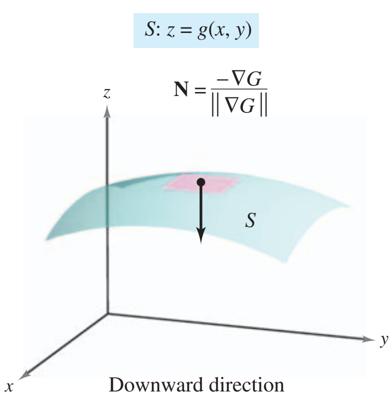
$$\nabla G(x, y, z) = \mathbf{i} - g_y(y, z)\mathbf{j} - g_z(y, z)\mathbf{k} \qquad G(x, y, z) = x - g(y, z)$$

to orient the surface.



S is oriented in an upward direction.

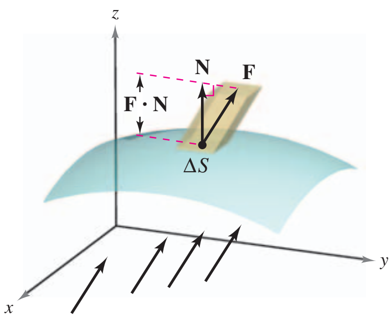
Rotatable Graph



S is oriented in a downward direction.

Rotatable Graph

Figure 15.50



The velocity field **F** indicates the direction of the fluid flow.
Figure 15.51

Rotatable Graph

Flux Integrals

One of the principal applications involving the vector form of a surface integral relates to the flow of a fluid through a surface S. Suppose an oriented surface S is submerged in a fluid having a continuous velocity field $\mathbf{F}$. Let ΔS be the area of a small patch of the surface S over which $\mathbf{F}$ is nearly constant. Then the amount of fluid crossing this region per unit of time is approximated by the volume of the column of height $\mathbf{F} \cdot \mathbf{N}$, as shown in Figure 15.51. That is,

$$\Delta V = (\text{height})(\text{area of base}) = (\mathbf{F} \cdot \mathbf{N})\Delta S.$$

Consequently, the volume of fluid crossing the surface S per unit of time (called the **flux of F across S**) is given by the surface integral in the following definition.

Definition of Flux Integral

Let $\mathbf{F}(x, y, z) = M\mathbf{i} + N\mathbf{j} + P\mathbf{k}$, where M, N, and P have continuous first partial derivatives on the surface S oriented by a unit normal vector $\mathbf{N}$. The **flux integral of F across S** is given by

$$\int_S\!\!\int \mathbf{F} \cdot \mathbf{N}\, dS.$$

Geometrically, a flux integral is the surface integral over S of the *normal component* of $\mathbf{F}$. If $\rho(x, y, z)$ is the density of the fluid at (x, y, z), the flux integral

$$\int_S\!\!\int \rho\, \mathbf{F} \cdot \mathbf{N}\, dS$$

represents the *mass* of the fluid flowing across S per unit of time.

To evaluate a flux integral for a surface given by $z = g(x, y)$, let

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

Then, $\mathbf{N}\, dS$ can be written as follows.

$$\begin{aligned}\mathbf{N}\, dS &= \frac{\nabla G(x, y, z)}{\|\nabla G(x, y, z)\|}\, dS \\ &= \frac{\nabla G(x, y, z)}{\sqrt{(g_x)^2 + (g_y)^2 + 1}} \sqrt{(g_x)^2 + (g_y)^2 + 1}\, dA \\ &= \nabla G(x, y, z)\, dA\end{aligned}$$

THEOREM 15.11 Evaluating a Flux Integral

Let S be an oriented surface given by $z = g(x, y)$ and let R be its projection onto the xy-plane.

$$\int_S\!\!\int \mathbf{F} \cdot \mathbf{N}\, dS = \int_R\!\!\int \mathbf{F} \cdot [-g_x(x, y)\mathbf{i} - g_y(x, y)\mathbf{j} + \mathbf{k}]\, dA \qquad \text{Oriented upward}$$

$$\int_S\!\!\int \mathbf{F} \cdot \mathbf{N}\, dS = \int_R\!\!\int \mathbf{F} \cdot [g_x(x, y)\mathbf{i} + g_y(x, y)\mathbf{j} - \mathbf{k}]\, dA \qquad \text{Oriented downward}$$

For the first integral, the surface is oriented upward, and for the second integral, the surface is oriented downward.

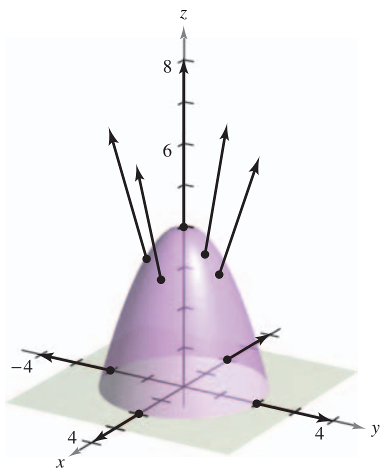


Figure 15.52

Rotatable Graph

EXAMPLE 5 Using a Flux Integral to Find the Rate of Mass Flow

Let S be the portion of the paraboloid

$$z = g(x, y) = 4 - x^2 - y^2$$

lying above the xy-plane, oriented by an upward unit normal vector, as shown in Figure 15.52. A fluid of constant density ρ is flowing through the surface S according to the vector field

$$\mathbf{F}(x, y, z) = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}.$$

Find the rate of mass flow through S.

Solution Begin by computing the partial derivatives of g.

$$g_x(x, y) = -2x$$

and

$$g_y(x, y) = -2y$$

The rate of mass flow through the surface S is

$$\begin{aligned}\int_S\int \rho\mathbf{F}\cdot\mathbf{N}\,dS &= \rho\int_R\int \mathbf{F}\cdot[-g_x(x, y)\mathbf{i} - g_y(x, y)\mathbf{j} + \mathbf{k}]\,dA\\ &= \rho\int_R\int [x\mathbf{i} + y\mathbf{j} + (4 - x^2 - y^2)\mathbf{k}]\cdot(2x\mathbf{i} + 2y\mathbf{j} + \mathbf{k})\,dA\\ &= \rho\int_R\int [2x^2 + 2y^2 + (4 - x^2 - y^2)]\,dA\\ &= \rho\int_R\int (4 + x^2 + y^2)\,dA\\ &= \rho\int_0^{2\pi}\int_0^2 (4 + r^2)r\,dr\,d\theta \qquad \text{Polar coordinates}\\ &= \rho\int_0^{2\pi} 12\,d\theta\\ &= 24\pi\rho.\end{aligned}$$

Try It Exploration A

For an oriented surface S given by the vector-valued function

$$\mathbf{r}(u, v) = x(u, v)\mathbf{i} + y(u, v)\mathbf{j} + z(u, v)\mathbf{k} \qquad \text{Parametric surface}$$

defined over a region D in the uv-plane, you can define the flux integral of $\mathbf{F}$ across S as

$$\begin{aligned}\int_S\int \mathbf{F}\cdot\mathbf{N}\,dS &= \int_D\int \mathbf{F}\cdot\left(\frac{\mathbf{r}_u\times\mathbf{r}_v}{\|\mathbf{r}_u\times\mathbf{r}_v\|}\right)\|\mathbf{r}_u\times\mathbf{r}_v\|\,dA\\ &= \int_D\int \mathbf{F}\cdot(\mathbf{r}_u\times\mathbf{r}_v)\,dA.\end{aligned}$$

Note the similarity of this integral to the line integral

$$\int_C \mathbf{F}\cdot d\mathbf{r} = \int_C \mathbf{F}\cdot\mathbf{T}\,ds.$$

A summary of formulas for line and surface integrals is presented on page 1117.

EXAMPLE 6 Finding the Flux of an Inverse Square Field

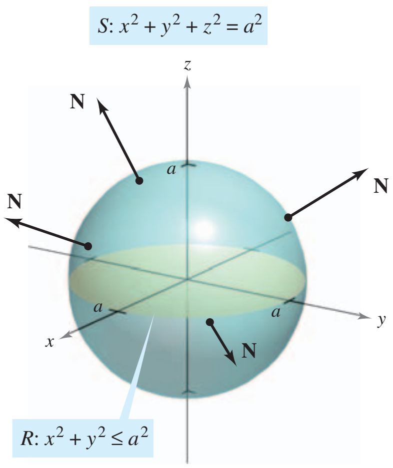


Figure 15.53

Rotatable Graph

Find the flux over the sphere S given by

$$x^2 + y^2 + z^2 = a^2 \qquad \text{Sphere } S$$

where $\mathbf{F}$ is an inverse square field given by

$$\mathbf{F}(x, y, z) = \frac{kq}{\|\mathbf{r}\|^2}\frac{\mathbf{r}}{\|\mathbf{r}\|} = \frac{kq\mathbf{r}}{\|\mathbf{r}\|^3} \qquad \text{Inverse square field } \mathbf{F}$$

and $\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$. Assume S is oriented outward, as shown in Figure 15.53.

Solution The sphere is given by

$$\begin{aligned}\mathbf{r}(u, v) &= x(u, v)\mathbf{i} + y(u, v)\mathbf{j} + z(u, v)\mathbf{k}\\ &= a \sin u \cos v\mathbf{i} + a \sin u \sin v\mathbf{j} + a \cos u\mathbf{k}\end{aligned}$$

where $0 \le u \le \pi$ and $0 \le v \le 2\pi$. The partial derivatives of $\mathbf{r}$ are

$$\mathbf{r}_u(u, v) = a \cos u \cos v\,\mathbf{i} + a \cos u \sin v\mathbf{j} - a \sin u\mathbf{k}$$

and

$$\mathbf{r}_v(u, v) = -a \sin u \sin v\mathbf{i} + a \sin u \cos v\mathbf{j}$$

which implies that the normal vector $\mathbf{r}_u \times \mathbf{r}_v$ is

$$\begin{aligned}\mathbf{r}_u \times \mathbf{r}_v &= \begin{vmatrix}\mathbf{i} & \mathbf{j} & \mathbf{k}\\ a \cos u \cos v & a \cos u \sin v & -a \sin u\\ -a \sin u \sin v & a \sin u \cos v & 0\end{vmatrix}\\ &= a^2(\sin^2 u \cos v\mathbf{i} + \sin^2 u \sin v\mathbf{j} + \sin u \cos u\mathbf{k}).\end{aligned}$$

Now, using

$$\begin{aligned}\mathbf{F}(x, y, z) &= \frac{kq\mathbf{r}}{\|\mathbf{r}\|^3}\\ &= kq\frac{x\mathbf{i} + y\mathbf{j} + z\mathbf{k}}{\|x\mathbf{i} + y\mathbf{j} + z\mathbf{k}\|^3}\\ &= \frac{kq}{a^3}(a \sin u \cos v\mathbf{i} + a \sin u \sin v\mathbf{j} + a \cos u\mathbf{k})\end{aligned}$$

it follows that

$$\begin{aligned}\mathbf{F} \cdot (\mathbf{r}_u \times \mathbf{r}_v) &= \frac{kq}{a^3}[(a \sin u \cos v\mathbf{i} + a \sin u \sin v\mathbf{j} + a \cos u\mathbf{k}) \cdot\\ &\qquad a^2(\sin^2 u \cos v\mathbf{i} + \sin^2 u \sin v\mathbf{j} + \sin u \cos u\mathbf{k})]\\ &= kq(\sin^3 u \cos^2 v + \sin^3 u \sin^2 v + \sin u \cos^2 u)\\ &= kq \sin u.\end{aligned}$$

Finally, the flux over the sphere S is given by

$$\begin{aligned}\int\!\!\int_S \mathbf{F} \cdot \mathbf{N}\, dS &= \int\!\!\int_D (kq \sin u)\, dA\\ &= \int_0^{2\pi}\!\!\int_0^{\pi} kq \sin u\, du\, dv\\ &= 4\pi kq.\end{aligned}$$

Try It

Exploration A

The result in Example 6 shows that the flux across a sphere S in an inverse square field is independent of the radius of S. In particular, if $\mathbf{E}$ is an electric field, the result in Example 6, along with Coulomb's Law, yields one of the basic laws of electrostatics, known as **Gauss's Law:**

$$\iint_S \mathbf{E} \cdot \mathbf{N}\, dS = 4\pi kq \qquad \text{Gauss's Law}$$

where q is a point charge located at the center of the sphere and k is the Coulomb constant. Gauss's Law is valid for more general closed surfaces that enclose the origin, and relates the flux out of the surface to the total charge q inside the surface.

This section concludes with a summary of different forms of line integrals and surface integrals.

Summary of Line and Surface Integrals

Line Integrals

$$ds = \|\mathbf{r}'(t)\|\, dt$$
$$= \sqrt{[x'(t)]^2 + [y'(t)]^2 + [z'(t)]^2}\, dt$$

$$\int_C f(x, y, z)\, ds = \int_a^b f(x(t), y(t), z(t))\, ds \qquad \text{Scalar form}$$

$$\int_C \mathbf{F} \cdot d\mathbf{r} = \int_C \mathbf{F} \cdot \mathbf{T}\, ds$$
$$= \int_a^b \mathbf{F}(x(t), y(t), z(t)) \cdot \mathbf{r}'(t)\, dt \qquad \text{Vector form}$$

Surface Integrals $[z = g(x, y)]$

$$dS = \sqrt{1 + [g_x(x, y)]^2 + [g_y(x, y)]^2}\, dA$$

$$\iint_S f(x, y, z)\, dS = \iint_R f(x, y, g(x, y))\sqrt{1 + [g_x(x, y)]^2 + [g_y(x, y)]^2}\, dA \qquad \text{Scalar form}$$

$$\iint_S \mathbf{F} \cdot \mathbf{N}\, dS = \iint_R \mathbf{F} \cdot [-g_x(x, y)\mathbf{i} - g_y(x, y)\mathbf{j} + \mathbf{k}]\, dA \qquad \text{Vector form (upward normal)}$$

Surface Integrals (*parametric form*)

$$dS = \|\mathbf{r}_u(u, v) \times \mathbf{r}_v(u, v)\|\, dA$$

$$\iint_S f(x, y, z)\, dS = \iint_D f(x(u, v), y(u, v), z(u, v))\, dS \qquad \text{Scalar form}$$

$$\iint_S \mathbf{F} \cdot \mathbf{N}\, dS = \iint_D \mathbf{F} \cdot (\mathbf{r}_u \times \mathbf{r}_v)\, dA \qquad \text{Vector form}$$

Section 15.7 Divergence Theorem

- **Understand and use the Divergence Theorem.**
- **Use the Divergence Theorem to calculate flux.**

CARL FRIEDRICH GAUSS (1777–1855)

The **Divergence Theorem** is also called **Gauss's Theorem,** after the famous German mathematician Carl Friedrich Gauss. Gauss is recognized, with Newton and Archimedes, as one of the three greatest mathematicians in history. One of his many contributions to mathematics was made at the age of 22, when, as part of his doctoral dissertation, he proved the *Fundamental Theorem of Algebra*.

MathBio

Divergence Theorem

Recall from Section 15.4 that an alternative form of Green's Theorem is

$$\int_C \mathbf{F}\cdot\mathbf{N}\,ds = \int\!\!\int_R \left(\frac{\partial M}{\partial x} + \frac{\partial N}{\partial y}\right) dA$$
$$= \int\!\!\int_R \operatorname{div}\mathbf{F}\,dA.$$

In an analogous way, the **Divergence Theorem** gives the relationship between a triple integral over a solid region Q and a surface integral over the surface of Q. In the statement of the theorem, the surface S is **closed** in the sense that it forms the complete boundary of the solid Q. Regions bounded by spheres, ellipsoids, cubes, tetrahedrons, or combinations of these surfaces are typical examples of closed surfaces. Assume that Q is a solid region on which a triple integral can be evaluated, and that the closed surface S is oriented by *outward* unit normal vectors, as shown in Figure 15.54. With these restrictions on S and Q, the Divergence Theorem is as follows.

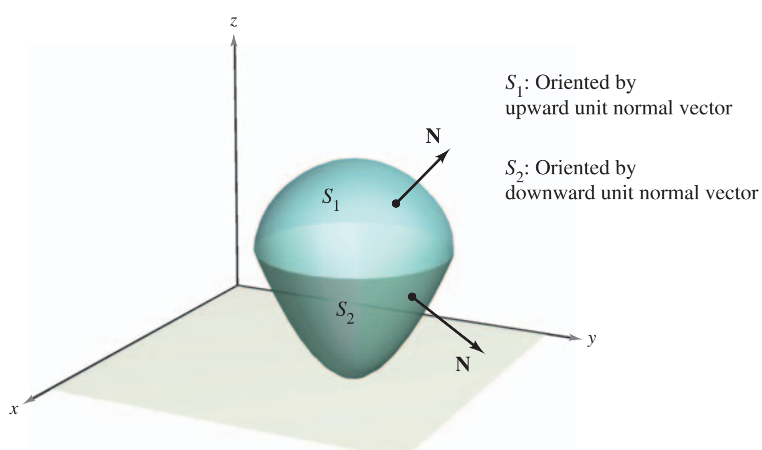


Figure 15.54

Rotatable Graph

THEOREM 15.12 The Divergence Theorem

Let Q be a solid region bounded by a closed surface S oriented by a unit normal vector directed outward from Q. If $\mathbf{F}$ is a vector field whose component functions have continuous partial derivatives in Q, then

$$\int\!\!\int_S \mathbf{F}\cdot\mathbf{N}\,dS = \int\!\!\int\!\!\int_Q \operatorname{div}\mathbf{F}\,dV.$$

NOTE As noted at the left above, the Divergence Theorem is sometimes called Gauss's Theorem. It is also sometimes called Ostrogradsky's Theorem, after the Russian mathematician Michel Ostrogradsky (1801–1861).

Proof If you let $\mathbf{F}(x, y, z) = M\mathbf{i} + N\mathbf{j} + P\mathbf{k}$, the theorem takes the form

$$\int_S\int \mathbf{F} \cdot \mathbf{N}\, dS = \int_S\int (M\mathbf{i} \cdot \mathbf{N} + N\mathbf{j} \cdot \mathbf{N} + P\mathbf{k} \cdot \mathbf{N})\, dS$$

$$= \iiint_Q \left(\frac{\partial M}{\partial x} + \frac{\partial N}{\partial y} + \frac{\partial P}{\partial z}\right) dV.$$

You can prove this by verifying that the following three equations are valid.

$$\int_S\int M\mathbf{i} \cdot \mathbf{N}\, dS = \iiint_Q \frac{\partial M}{\partial x}\, dV$$

$$\int_S\int N\mathbf{j} \cdot \mathbf{N}\, dS = \iiint_Q \frac{\partial N}{\partial y}\, dV$$

$$\int_S\int P\mathbf{k} \cdot \mathbf{N}\, dS = \iiint_Q \frac{\partial P}{\partial z}\, dV$$

Because the verifications of the three equations are similar, only the third is discussed. Restrict the proof to a **simple solid** region with upper surface

$z = g_2(x, y)$ Upper surface

and lower surface

$z = g_1(x, y)$ Lower surface

whose projections onto the xy-plane coincide and form region R. If Q has a lateral surface like S_3 in Figure 15.55, then a normal vector is horizontal, which implies that $P\mathbf{k} \cdot \mathbf{N} = 0$. Consequently, you have

$$\int_S\int P\mathbf{k} \cdot \mathbf{N}\, dS = \int_{S_1}\int P\mathbf{k} \cdot \mathbf{N}\, dS + \int_{S_2}\int P\mathbf{k} \cdot \mathbf{N}\, dS + 0.$$

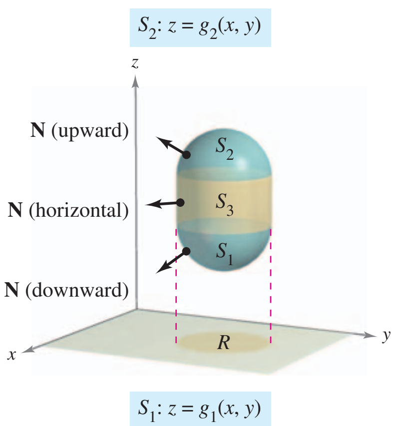


Figure 15.55

Rotatable Graph

On the upper surface S_2, the outward normal vector is upward, whereas on the lower surface S_1, the outward normal vector is downward. So, by Theorem 15.11, you have the following.

$$\int_{S_1}\int P\mathbf{k} \cdot \mathbf{N}\, dS = \int_R\int P(x, y, g_1(x, y))\mathbf{k} \cdot \left(\frac{\partial g_1}{\partial x}\mathbf{i} + \frac{\partial g_1}{\partial y}\mathbf{j} - \mathbf{k}\right) dA$$

$$= -\int_R\int P(x, y, g_1(x, y))\, dA$$

$$\int_{S_2}\int P\mathbf{k} \cdot \mathbf{N}\, dS = \int_R\int P(x, y, g_2(x, y))\mathbf{k} \cdot \left(-\frac{\partial g_2}{\partial x}\mathbf{i} - \frac{\partial g_2}{\partial y}\mathbf{j} + \mathbf{k}\right) dA$$

$$= \int_R\int P(x, y, g_2(x, y))\, dA$$

Adding these results, you obtain

$$\int_S\int P\mathbf{k} \cdot \mathbf{N}\, dS = \int_R\int [P(x, y, g_2(x, y)) - P(x, y, g_1(x, y))]\, dA$$

$$= \int_R\int \left[\int_{g_1(x, y)}^{g_2(x, y)} \frac{\partial P}{\partial z}\, dz\right] dA$$

$$= \iiint_Q \frac{\partial P}{\partial z}\, dV.$$

EXAMPLE 1 Using the Divergence Theorem

Let Q be the solid region bounded by the coordinate planes and the plane $2x + 2y + z = 6$, and let $\mathbf{F} = x\mathbf{i} + y^2\mathbf{j} + z\mathbf{k}$. Find

$$\int_S\!\!\int \mathbf{F} \cdot \mathbf{N}\, dS$$

where S is the surface of Q.

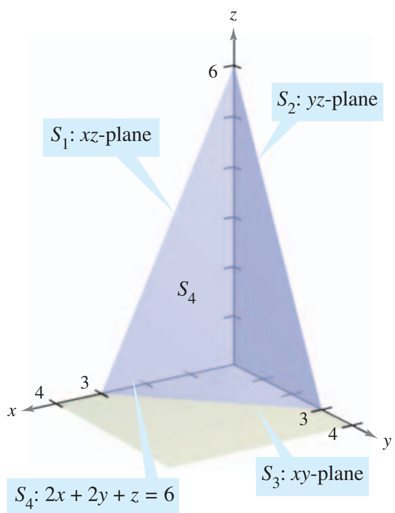


Figure 15.56

Rotatable Graph

Solution From Figure 15.56, you can see that Q is bounded by four subsurfaces. So, you would need four *surface integrals* to evaluate

$$\int_S\!\!\int \mathbf{F} \cdot \mathbf{N}\, dS.$$

However, by the Divergence Theorem, you need only one triple integral. Because

$$\begin{aligned}\text{div}\,\mathbf{F} &= \frac{\partial M}{\partial x} + \frac{\partial N}{\partial y} + \frac{\partial P}{\partial z}\\ &= 1 + 2y + 1\\ &= 2 + 2y\end{aligned}$$

you have

$$\begin{aligned}\int_S\!\!\int \mathbf{F} \cdot \mathbf{N}\, dS &= \iiint_Q \text{div}\,\mathbf{F}\, dV\\ &= \int_0^3\int_0^{3-y}\int_0^{6-2x-2y} (2 + 2y)\, dz\, dx\, dy\\ &= \int_0^3\int_0^{3-y} \Big[(2z + 2yz)\Big]_0^{6-2x-2y} dx\, dy\\ &= \int_0^3\int_0^{3-y} (12 - 4x + 8y - 4xy - 4y^2)\, dx\, dy\\ &= \int_0^3 \Big[12x - 2x^2 + 8xy - 2x^2y - 4xy^2\Big]_0^{3-y} dy\\ &= \int_0^3 (18 + 6y - 10y^2 + 2y^3)\, dy\\ &= \Big[18y + 3y^2 - \frac{10y^3}{3} + \frac{y^4}{2}\Big]_0^3\\ &= \frac{63}{2}.\end{aligned}$$

Try It Exploration A

TECHNOLOGY If you have access to a computer algebra system that can evaluate triple-iterated integrals, use it to verify the result in Example 1. When you are using such a utility, note that the first step is to convert the triple integral to an iterated integral—this step must be done by hand. To give yourself some practice with this important step, find the limits of integration for the following iterated integrals. Then use a computer to verify that the value is the same as that obtained in Example 1.

$$\int_?^?\int_?^?\int_?^? (2 + 2y)\, dy\, dz\, dx, \qquad \int_?^?\int_?^?\int_?^? (2 + 2y)\, dx\, dy\, dz$$

EXAMPLE 2 Verifying the Divergence Theorem

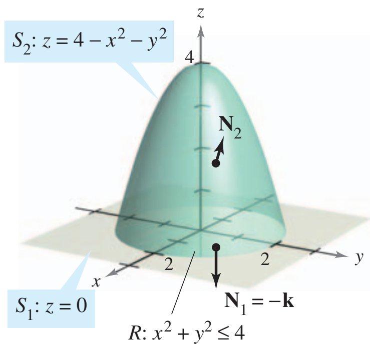


Figure 15.57

Rotatable Graph

Let Q be the solid region between the paraboloid

$$z = 4 - x^2 - y^2$$

and the xy-plane. Verify the Divergence Theorem for

$$\mathbf{F}(x, y, z) = 2z\mathbf{i} + x\mathbf{j} + y^2\mathbf{k}.$$

Solution From Figure 15.57 you can see that the outward normal vector for the surface S_1 is $\mathbf{N}_1 = -\mathbf{k}$, whereas the outward normal vector for the surface S_2 is

$$\mathbf{N}_2 = \frac{2x\mathbf{i} + 2y\mathbf{j} + \mathbf{k}}{\sqrt{4x^2 + 4y^2 + 1}}.$$

So, by Theorem 15.11, you have

$$\begin{aligned}
\int_S\!\!\int \mathbf{F}\cdot\mathbf{N}\,dS
&= \int_{S_1}\!\!\int \mathbf{F}\cdot\mathbf{N}_1\,dS + \int_{S_2}\!\!\int \mathbf{F}\cdot\mathbf{N}_2\,dS\\
&= \int_{S_1}\!\!\int \mathbf{F}\cdot(-\mathbf{k})\,dS + \int_{S_2}\!\!\int \mathbf{F}\cdot(2x\mathbf{i} + 2y\mathbf{j} + \mathbf{k})\,dS\\
&= \int_R\!\!\int -y^2\,dA + \int_R\!\!\int (4xz + 2xy + y^2)\,dA\\
&= -\int_{-2}^{2}\int_{-\sqrt{4-y^2}}^{\sqrt{4-y^2}} y^2\,dx\,dy + \int_{-2}^{2}\int_{-\sqrt{4-y^2}}^{\sqrt{4-y^2}} (4xz + 2xy + y^2)\,dx\,dy\\
&= \int_{-2}^{2}\int_{-\sqrt{4-y^2}}^{\sqrt{4-y^2}} (4xz + 2xy)\,dx\,dy\\
&= \int_{-2}^{2}\int_{-\sqrt{4-y^2}}^{\sqrt{4-y^2}} [4x(4 - x^2 - y^2) + 2xy]\,dx\,dy\\
&= \int_{-2}^{2}\int_{-\sqrt{4-y^2}}^{\sqrt{4-y^2}} (16x - 4x^3 - 4xy^2 + 2xy)\,dx\,dy\\
&= \int_{-2}^{2}\Big[8x^2 - x^4 - 2x^2y^2 + x^2y\Big]_{-\sqrt{4-y^2}}^{\sqrt{4-y^2}}\,dy\\
&= \int_{-2}^{2} 0\,dy\\
&= 0.
\end{aligned}$$

On the other hand, because

$$\text{div}\,\mathbf{F} = \frac{\partial}{\partial x}[2z] + \frac{\partial}{\partial y}[x] + \frac{\partial}{\partial z}[y^2] = 0 + 0 + 0 = 0$$

you can apply the Divergence Theorem to obtain the equivalent result

$$\begin{aligned}
\int_S\!\!\int \mathbf{F}\cdot\mathbf{N}\,dS &= \iiint_Q \text{div}\,\mathbf{F}\,dV\\
&= \iiint_Q 0\,dV = 0.
\end{aligned}$$

Try It

Exploration A

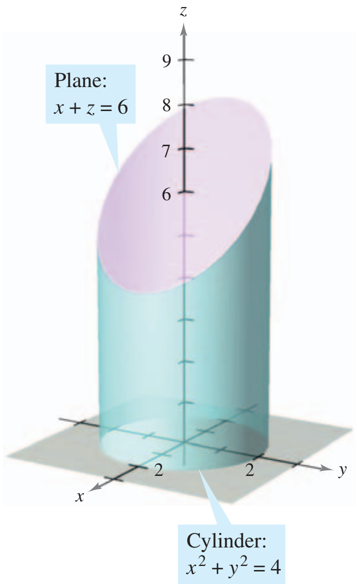


Figure 15.58

Rotatable Graph

EXAMPLE 3 Using the Divergence Theorem

Let Q be the solid bounded by the cylinder $x^2 + y^2 = 4$, the plane $x + z = 6$, and the xy-plane, as shown in Figure 15.58. Find

$$\int_S\int \mathbf{F} \cdot \mathbf{N}\, dS$$

where S is the surface of Q and

$$\mathbf{F}(x, y, z) = (x^2 + \sin z)\mathbf{i} + (xy + \cos z)\mathbf{j} + e^y\mathbf{k}.$$

Solution Direct evaluation of this surface integral would be difficult. However, by the Divergence Theorem, you can evaluate the integral as follows.

$$\begin{aligned}
\int_S\int \mathbf{F} \cdot \mathbf{N}\, dS &= \iiint_Q \text{div}\,\mathbf{F}\, dV \\
&= \iiint_Q (2x + x + 0)\, dV \\
&= \iiint_Q 3x\, dV \\
&= \int_0^{2\pi}\int_0^2\int_0^{6-r\cos\theta} (3r\cos\theta)r\, dz\, dr\, d\theta \\
&= \int_0^{2\pi}\int_0^2 (18r^2\cos\theta - 3r^3\cos^2\theta)\, dr\, d\theta \\
&= \int_0^{2\pi} (48\cos\theta - 12\cos^2\theta)d\theta \\
&= \left[48\sin\theta - 6\left(\theta + \frac{1}{2}\sin 2\theta\right)\right]_0^{2\pi} \\
&= -12\pi
\end{aligned}$$

Notice that cylindrical coordinates with $x = r\cos\theta$ and $dV = r\, dz\, dr\, d\theta$ were used to evaluate the triple integral.

Try It Exploration A

Even though the Divergence Theorem was stated for a simple solid region Q bounded by a closed surface, the theorem is also valid for regions that are the finite unions of simple solid regions. For example, let Q be the solid bounded by the closed surfaces S_1 and S_2, as shown in Figure 15.59. To apply the Divergence Theorem to this solid, let $S = S_1 \cup S_2$. The normal vector $\mathbf{N}$ to S is given by $-\mathbf{N}_1$ on S_1 and by $\mathbf{N}_2$ on S_2. So, you can write

$$\begin{aligned}
\iiint_Q \text{div}\,\mathbf{F}\, dV &= \int_S\int \mathbf{F} \cdot \mathbf{N}\, dS \\
&= \int_{S_1}\int \mathbf{F} \cdot (-\mathbf{N}_1)\, dS + \int_{S_2}\int \mathbf{F} \cdot \mathbf{N}_2\, dS \\
&= -\int_{S_1}\int \mathbf{F} \cdot \mathbf{N}_1\, dS + \int_{S_2}\int \mathbf{F} \cdot \mathbf{N}_2\, dS.
\end{aligned}$$

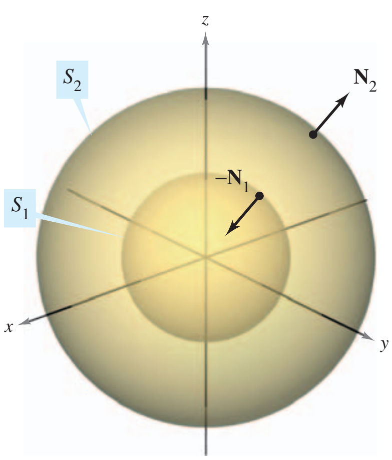


Figure 15.59

Rotatable Graph

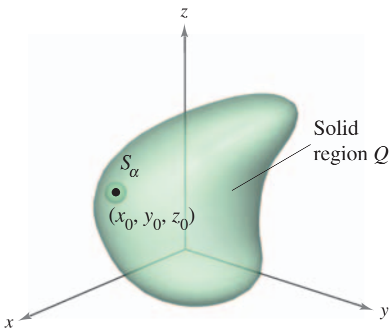
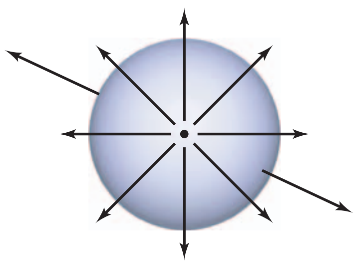


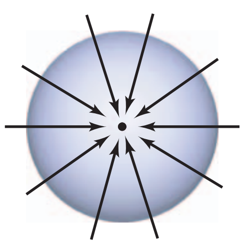
Figure 15.60

Rotatable Graph



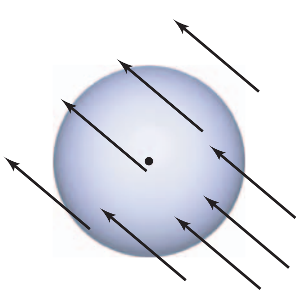
Rotatable Graph

(a) Source: div $\mathbf{F} > 0$



Rotatable Graph

(b) Sink: div $\mathbf{F} < 0$



Rotatable Graph

(c) Incompressible: div $\mathbf{F} = 0$

Figure 15.61

Flux and the Divergence Theorem

To help understand the Divergence Theorem, consider the two sides of the equation

$$\int_S\!\!\int \mathbf{F}\cdot\mathbf{N}\,dS = \iiint_Q \text{div}\,\mathbf{F}\,dV.$$

You know from Section 15.6 that the flux integral on the left determines the total fluid flow across the surface S per unit of time. This can be approximated by summing the fluid flow across small patches of the surface. The triple integral on the right measures this same fluid flow across S, but from a very different perspective—namely, by calculating the flow of fluid into (or out of) small *cubes* of volume ΔV_i. The flux of the ith cube is approximately

$$\text{Flux of } i\text{th cube} \approx \text{div}\,\mathbf{F}(x_i, y_i, z_i)\,\Delta V_i$$

for some point (x_i, y_i, z_i) in the ith cube. Note that for a cube in the interior of Q, the gain (or loss) of fluid through any one of its six sides is offset by a corresponding loss (or gain) through one of the sides of an adjacent cube. After summing over all the cubes in Q, the only fluid flow that is not canceled by adjoining cubes is that on the outside edges of the cubes on the boundary. So, the sum

$$\sum_{i=1}^{n} \text{div}\,\mathbf{F}(x_i, y_i, z_i)\,\Delta V_i$$

approximates the total flux into (or out of) Q, and therefore through the surface S.

To see what is meant by the divergence of $\mathbf{F}$ at a point, consider ΔV_α to be the volume of a small sphere S_α of radius α and center (x_0, y_0, z_0), contained in region Q, as shown in Figure 15.60. Applying the Divergence Theorem to S_α produces

$$\begin{aligned}\text{Flux of } \mathbf{F} \text{ across } S_\alpha &= \iiint_{Q_\alpha} \text{div}\,\mathbf{F}\,dV\\ &\approx \text{div}\,\mathbf{F}(x_0, y_0, z_0)\,\Delta V_\alpha\end{aligned}$$

where Q_α is the interior of S_α. Consequently, you have

$$\text{div}\,\mathbf{F}(x_0, y_0, z_0) \approx \frac{\text{flux of } \mathbf{F} \text{ across } S_\alpha}{\Delta V_\alpha}$$

and, by taking the limit as $\alpha \to 0$, you obtain the divergence of $\mathbf{F}$ at the point (x_0, y_0, z_0).

$$\begin{aligned}\text{div}\,\mathbf{F}(x_0, y_0, z_0) &= \lim_{\alpha\to 0} \frac{\text{flux of } \mathbf{F} \text{ across } S_\alpha}{\Delta V_\alpha}\\ &= \text{flux per unit volume at } (x_0, y_0, z_0)\end{aligned}$$

The point (x_0, y_0, z_0) in a vector field is classified as a source, a sink, or incompressible, as follows.

1. **Source,** if div $\mathbf{F} > 0$ — See Figure 15.61(a).
2. **Sink,** if div $\mathbf{F} < 0$ — See Figure 15.61(b).
3. **Incompressible,** if div $\mathbf{F} = 0$ — See Figure 15.61(c).

NOTE In hydrodynamics, a *source* is a point at which additional fluid is considered as being introduced to the region occupied by the fluid. A *sink* is a point at which fluid is considered as being removed.

EXAMPLE 4 Calculating Flux by the Divergence Theorem

Let Q be the region bounded by the sphere $x^2 + y^2 + z^2 = 4$. Find the outward flux of the vector field $\mathbf{F}(x, y, z) = 2x^3\mathbf{i} + 2y^3\mathbf{j} + 2z^3\mathbf{k}$ through the sphere.

Solution By the Divergence Theorem, you have

$$\begin{aligned}
\text{Flux across } S &= \int\!\!\int_S \mathbf{F}\cdot\mathbf{N}\,dS = \int\!\!\int\!\!\int_Q \operatorname{div}\mathbf{F}\,dV \\
&= \int\!\!\int\!\!\int_Q 6(x^2 + y^2 + z^2)\,dV \\
&= 6\int_0^2\int_0^{\pi}\int_0^{2\pi} \rho^4 \sin\phi\,d\theta\,d\phi\,d\rho \qquad \text{Spherical coordinates} \\
&= 6\int_0^2\int_0^{\pi} 2\pi\rho^4 \sin\phi\,d\phi\,d\rho \\
&= 12\pi\int_0^2 2\rho^4\,d\rho \\
&= 24\pi\left(\frac{32}{5}\right) \\
&= \frac{768\pi}{5}.
\end{aligned}$$

Try It | Exploration A | Exploration B | Open Exploration

Section 15.8 Stokes's Theorem

- **Understand and use Stokes's Theorem.**
- **Use curl to analyze the motion of a rotating liquid.**

George Gabriel Stokes (1819–1903)

Stokes became a Lucasian professor of mathematics at Cambridge in 1849. Five years later, he published the theorem that bears his name as a prize examination question there.

MathBio

Stokes's Theorem

A second higher-dimension analog of Green's Theorem is called **Stokes's Theorem,** after the English mathematical physicist George Gabriel Stokes. Stokes was part of a group of English mathematical physicists referred to as the Cambridge School, which included William Thomson (Lord Kelvin) and James Clerk Maxwell. In addition to making contributions to physics, Stokes worked with infinite series and differential equations, as well as with the integration results presented in this section.

Stokes's Theorem gives the relationship between a surface integral over an oriented surface S and a line integral along a closed space curve C forming the boundary of S, as shown in Figure 15.62. The positive direction along C is counterclockwise relative to the normal vector $\mathbf{N}$. That is, if you imagine grasping the normal vector $\mathbf{N}$ with your right hand, with your thumb pointing in the direction of $\mathbf{N}$, your fingers will point in the positive direction C, as shown in Figure 15.63.

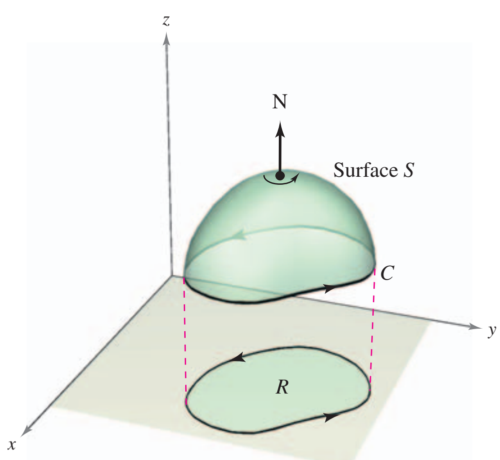
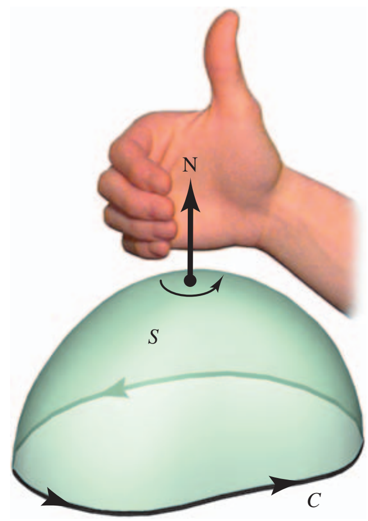


Figure 15.62

Rotatable Graph



Direction along C is counterclockwise relative to $\mathbf{N}$.
Figure 15.63

Rotatable Graph

THEOREM 15.13 Stokes's Theorem

Let S be an oriented surface with unit normal vector $\mathbf{N}$, bounded by a piecewise smooth simple closed curve C with a positive orientation. If $\mathbf{F}$ is a vector field whose component functions have continuous partial derivatives on an open region containing S and C, then

$$\int_C \mathbf{F} \cdot d\mathbf{r} = \int_S\int (\mathbf{curl\ F}) \cdot \mathbf{N}\, dS.$$

NOTE The line integral may be written in the differential form $\int_C M\,dx + N\,dy + P\,dz$ or in the vector form $\int_C \mathbf{F} \cdot \mathbf{T}\, ds$.

EXAMPLE 1 Using Stokes's Theorem

Let C be the oriented triangle lying in the plane $2x + 2y + z = 6$, as shown in Figure 15.64. Evaluate

$$\int_C \mathbf{F} \cdot d\mathbf{r}$$

where $\mathbf{F}(x, y, z) = -y^2\mathbf{i} + z\mathbf{j} + x\mathbf{k}$.

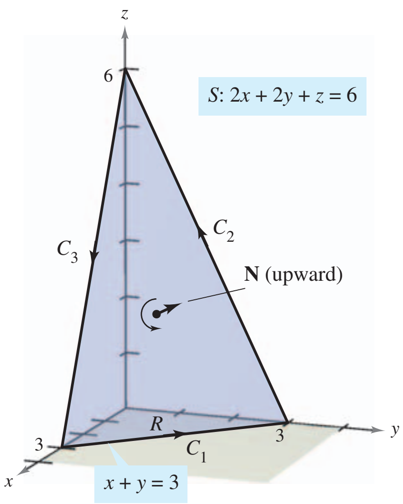


Figure 15.64

Rotatable Graph

Solution Using Stokes's Theorem, begin by finding the curl of $\mathbf{F}$.

$$\textbf{curl F} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ -y^2 & z & x \end{vmatrix} = -\mathbf{i} - \mathbf{j} + 2y\mathbf{k}$$

Considering $z = 6 - 2x - 2y = g(x, y)$, you can use Theorem 15.11 for an upward normal vector to obtain

$$\begin{aligned}
\int_C \mathbf{F} \cdot d\mathbf{r} &= \int_S\int (\textbf{curl F}) \cdot \mathbf{N}\, dS \\
&= \int_R\int (-\mathbf{i} - \mathbf{j} + 2y\mathbf{k}) \cdot [-g_x(x, y)\mathbf{i} - g_y(x, y)\mathbf{j} + \mathbf{k}]\, dA \\
&= \int_R\int (-\mathbf{i} - \mathbf{j} + 2y\mathbf{k}) \cdot (2\mathbf{i} + 2\mathbf{j} + \mathbf{k})\, dA \\
&= \int_0^3\int_0^{3-y} (2y - 4)\, dx\, dy \\
&= \int_0^3 (-2y^2 + 10y - 12)\, dy \\
&= \left[-\frac{2y^3}{3} + 5y^2 - 12y\right]_0^3 \\
&= -9.
\end{aligned}$$

Try It | Exploration A

Try evaluating the line integral in Example 1 directly, *without* using Stokes's Theorem. One way to do this would be to consider C as the union of C_1, C_2, and C_3, as follows.

$$\begin{aligned}
&C_1\text{: } \mathbf{r}_1(t) = (3 - t)\mathbf{i} + t\mathbf{j}, \quad 0 \le t \le 3 \\
&C_2\text{: } \mathbf{r}_2(t) = (6 - t)\mathbf{j} + (2t - 6)\mathbf{k}, \quad 3 \le t \le 6 \\
&C_3\text{: } \mathbf{r}_3(t) = (t - 6)\mathbf{i} + (18 - 2t)\mathbf{k}, \quad 6 \le t \le 9
\end{aligned}$$

The value of the line integral is

$$\begin{aligned}
\int_C \mathbf{F} \cdot d\mathbf{r} &= \int_{C_1} \mathbf{F} \cdot \mathbf{r}_1'(t)\, dt + \int_{C_2} \mathbf{F} \cdot \mathbf{r}_2'(t)\, dt + \int_{C_3} \mathbf{F} \cdot \mathbf{r}_3'(t)\, dt \\
&= \int_0^3 t^2\, dt + \int_3^6 (-2t + 6)\, dt + \int_6^9 (-2t + 12)\, dt \\
&= 9 - 9 - 9 \\
&= -9.
\end{aligned}$$

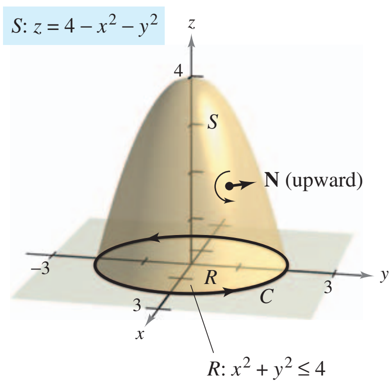


Figure 15.65

Rotatable Graph

EXAMPLE 2 Verifying Stokes's Theorem

Verify Stokes's Theorem for $\mathbf{F}(x, y, z) = 2z\mathbf{i} + x\mathbf{j} + y^2\mathbf{k}$, where S is the surface of the paraboloid $z = 4 - x^2 - y^2$ and C is the trace of S in the xy-plane, as shown in Figure 15.65.

Solution As a *surface integral*, you have $z = g(x, y) = 4 - x^2 - y^2$ and

$$\textbf{curl F} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ 2z & x & y^2 \end{vmatrix} = 2y\mathbf{i} + 2\mathbf{j} + \mathbf{k}.$$

By Theorem 15.11 for an upward normal vector $\mathbf{N}$, you obtain

$$\begin{aligned}
\int_S\!\!\int (\textbf{curl F}) \cdot \mathbf{N}\, dS &= \int_R\!\!\int (2y\mathbf{i} + 2\mathbf{j} + \mathbf{k}) \cdot (2x\mathbf{i} + 2y\mathbf{j} + \mathbf{k})\, dA \\
&= \int_{-2}^{2} \int_{-\sqrt{4-y^2}}^{\sqrt{4-y^2}} (4xy + 4y + 1)\, dx\, dy \\
&= \int_{-2}^{2} \Big[2x^2y + (4y + 1)x\Big]_{-\sqrt{4-y^2}}^{\sqrt{4-y^2}} dy \\
&= \int_{-2}^{2} 2(4y + 1)\sqrt{4 - y^2}\, dy \\
&= \int_{-2}^{2} \left(8y\sqrt{4 - y^2} + 2\sqrt{4 - y^2}\right) dy \\
&= \left[-\frac{8}{3}(4 - y^2)^{3/2} + y\sqrt{4 - y^2} + 4 \arcsin \frac{y}{2}\right]_{-2}^{2} \\
&= 4\pi.
\end{aligned}$$

As a *line integral*, you can parametrize C by

$$\mathbf{r}(t) = 2 \cos t\,\mathbf{i} + 2 \sin t\,\mathbf{j} + 0\mathbf{k}, \quad 0 \le t \le 2\pi.$$

For $\mathbf{F}(x, y, z) = 2z\mathbf{i} + x\mathbf{j} + y^2\mathbf{k}$, you obtain

$$\begin{aligned}
\int_C \mathbf{F} \cdot d\mathbf{r} &= \int_C M\, dx + N\, dy + P\, dz \\
&= \int_C 2z\, dx + x\, dy + y^2\, dz \\
&= \int_0^{2\pi} [0 + 2 \cos t(2 \cos t) + 0]\, dt \\
&= \int_0^{2\pi} 4 \cos^2 t\, dt \\
&= 2\int_0^{2\pi} (1 + \cos 2t)\, dt \\
&= 2\left[t + \frac{1}{2} \sin 2t\right]_0^{2\pi} \\
&= 4\pi.
\end{aligned}$$

Try It | Exploration A | Open Exploration

Physical Interpretation of Curl

Stokes's Theorem provides insight into a physical interpretation of curl. In a vector field **F**, let S_α be a *small* circular disk of radius α, centered at (x, y, z) and with boundary C_α, as shown in Figure 15.66. At each point on the circle C_α, **F** has a normal component $\mathbf{F} \cdot \mathbf{N}$ and a tangential component $\mathbf{F} \cdot \mathbf{T}$. The more closely **F** and **T** are aligned, the greater the value of $\mathbf{F} \cdot \mathbf{T}$. So, a fluid tends to move along the circle rather than across it. Consequently, you say that the line integral around C_α measures the **circulation of F around C_α.** That is,

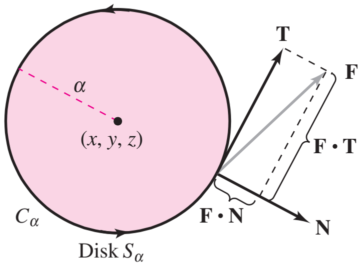


Figure 15.66

$$\int_{C_\alpha} \mathbf{F} \cdot \mathbf{T}\, ds = \text{circulation of } \mathbf{F} \text{ around } C_\alpha.$$

Now consider a small disk S_α to be centered at some point (x, y, z) on the surface S, as shown in Figure 15.67. On such a small disk, **curl F** is nearly constant, because it varies little from its value at (x, y, z). Moreover, $\textbf{curl F} \cdot \mathbf{N}$ is also nearly constant on S_α, because all unit normals to S_α are about the same. Consequently, Stokes's Theorem yields

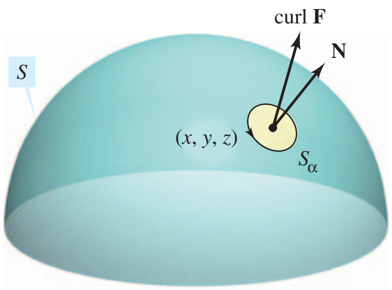


Figure 15.67

$$\begin{aligned}
\int_{C_\alpha} \mathbf{F} \cdot \mathbf{T}\, ds &= \int_{S_\alpha}\int (\textbf{curl F}) \cdot \mathbf{N}\, dS \\
&\approx (\textbf{curl F}) \cdot \mathbf{N} \int_{S_\alpha}\int dS \\
&\approx (\textbf{curl F}) \cdot \mathbf{N}(\pi\alpha^2).
\end{aligned}$$

So,

$$\begin{aligned}
(\textbf{curl F}) \cdot \mathbf{N} &\approx \frac{\displaystyle\int_{C_\alpha} \mathbf{F} \cdot \mathbf{T}\, ds}{\pi\alpha^2} \\
&= \frac{\text{circulation of } \mathbf{F} \text{ around } C_\alpha}{\text{area of disk } S_\alpha} \\
&= \text{rate of circulation.}
\end{aligned}$$

Assuming conditions are such that the approximation improves for smaller and smaller disks $(\alpha \to 0)$, it follows that

$$(\textbf{curl F}) \cdot \mathbf{N} = \lim_{\alpha \to 0} \frac{1}{\pi\alpha^2} \int_{C_\alpha} \mathbf{F} \cdot \mathbf{T}\, ds$$

which is referred to as the **rotation of F about N.** That is,

$$\textbf{curl F}(x, y, z) \cdot \mathbf{N} = \text{rotation of } \mathbf{F} \text{ about } \mathbf{N} \text{ at } (x, y, z).$$

In this case, the rotation of **F** is maximum when curl **F** and **N** have the same direction. Normally, this tendency to rotate will vary from point to point on the surface S, and Stokes's Theorem

$$\underbrace{\int_S\int (\textbf{curl F}) \cdot \mathbf{N}\, dS}_{\text{Surface integral}} = \underbrace{\int_C \mathbf{F} \cdot d\mathbf{r}}_{\text{Line integral}}$$

says that the collective measure of this *rotational* tendency taken over the entire surface S (surface integral) is equal to the tendency of a fluid to *circulate* around the boundary C (line integral).

EXAMPLE 3 An Application of Curl

A liquid is swirling around in a cylindrical container of radius 2, so that its motion is described by the velocity field

$$\mathbf{F}(x, y, z) = -y\sqrt{x^2 + y^2}\,\mathbf{i} + x\sqrt{x^2 + y^2}\,\mathbf{j}$$

as shown in Figure 15.68. Find

$$\int_S\int (\mathbf{curl\ F}) \cdot \mathbf{N}\, dS$$

where S is the upper surface of the cylindrical container.

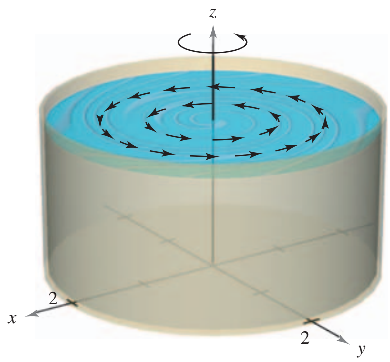


Figure 15.68

Rotatable Graph

Solution The curl of $\mathbf{F}$ is given by

$$\mathbf{curl\ F} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ -y\sqrt{x^2 + y^2} & x\sqrt{x^2 + y^2} & 0 \end{vmatrix} = 3\sqrt{x^2 + y^2}\,\mathbf{k}.$$

Letting $\mathbf{N} = \mathbf{k}$, you have

$$\begin{aligned}\int_S\int (\mathbf{curl\ F}) \cdot \mathbf{N}\, dS &= \int_R\int 3\sqrt{x^2 + y^2}\, dA \\ &= \int_0^{2\pi}\int_0^2 (3r)r\, dr\, d\theta \\ &= \int_0^{2\pi} r^3\Big]_0^2\, d\theta \\ &= \int_0^{2\pi} 8\, d\theta \\ &= 16\pi.\end{aligned}$$

Try It

Exploration A

NOTE If $\mathbf{curl\ F} = \mathbf{0}$ throughout region Q, the rotation of $\mathbf{F}$ about each unit normal $\mathbf{N}$ is 0. That is, $\mathbf{F}$ is irrotational. From earlier work, you know that this is a characteristic of conservative vector fields.

Summary of Integration Formulas

Fundamental Theorem of Calculus:

$$\int_a^b F'(x)\, dx = F(b) - F(a)$$

Fundamental Theorem of Line Integrals:

$$\int_C \mathbf{F} \cdot d\mathbf{r} = \int_C \nabla f \cdot d\mathbf{r} = f(x(b), y(b)) - f(x(a), y(a))$$

Green's Theorem:

$$\int_C M\, dx + N\, dy = \int_R\int \left(\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y}\right) dA = \int_C \mathbf{F} \cdot \mathbf{T}\, ds = \int_C \mathbf{F} \cdot d\mathbf{r} = \int_R\int (\mathbf{curl\ F}) \cdot \mathbf{k}\, dA$$

$$\int_C \mathbf{F} \cdot \mathbf{N}\, ds = \int_R\int \text{div}\,\mathbf{F}\, dA$$

Divergence Theorem:

$$\int_S\int \mathbf{F} \cdot \mathbf{N}\, dS = \iiint_Q \text{div}\,\mathbf{F}\, dV$$

Stokes's Theorem:

$$\int_C \mathbf{F} \cdot d\mathbf{r} = \int_S\int (\mathbf{curl\ F}) \cdot \mathbf{N}\, dS$$