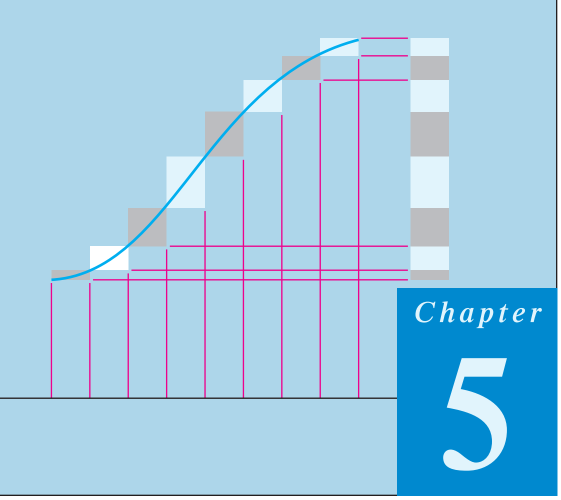


Chapter 5

INTEGRATION

OVERVIEW One of the great achievements of classical geometry was to obtain formulas for the areas and volumes of triangles, spheres, and cones. In this chapter we study a method to calculate the areas and volumes of these and other more general shapes. The method we develop, called *integration*, is a tool for calculating much more than areas and volumes. The *integral* has many applications in statistics, economics, the sciences, and engineering. It allows us to calculate quantities ranging from probabilities and averages to energy consumption and the forces against a dam's floodgates.

The idea behind integration is that we can effectively compute many quantities by breaking them into small pieces, and then summing the contributions from each small part. We develop the theory of the integral in the setting of area, where it most clearly reveals its nature. We begin with examples involving finite sums. These lead naturally to the question of what happens when more and more terms are summed. Passing to the limit, as the number of terms goes to infinity, then gives an integral. While integration and differentiation are closely connected, we will not see the roles of the derivative and antiderivative emerge until Section 5.4. The nature of their connection, contained in the Fundamental Theorem of Calculus, is one of the most important ideas in calculus.

5.1 Estimating with Finite Sums

This section shows how area, average values, and the distance traveled by an object over time can all be approximated by finite sums. Finite sums are the basis for defining the integral in Section 5.3.

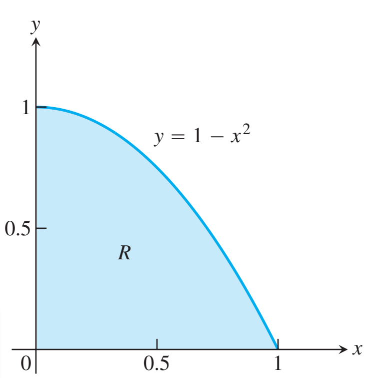


FIGURE 5.1 The area of the region R cannot be found by a simple geometry formula (Example 1).

Area

The area of a region with a curved boundary can be approximated by summing the areas of a collection of rectangles. Using more rectangles can increase the accuracy of the approximation.

EXAMPLE 1 Approximating Area

What is the area of the shaded region R that lies above the x-axis, below the graph of $y = 1 - x^2$, and between the vertical lines $x = 0$ and $x = 1$? (See Figure 5.1.) An architect might want to know this area to calculate the weight of a custom window with a shape described by R. Unfortunately, there is no simple geometric formula for calculating the areas of shapes having curved boundaries like the region R.

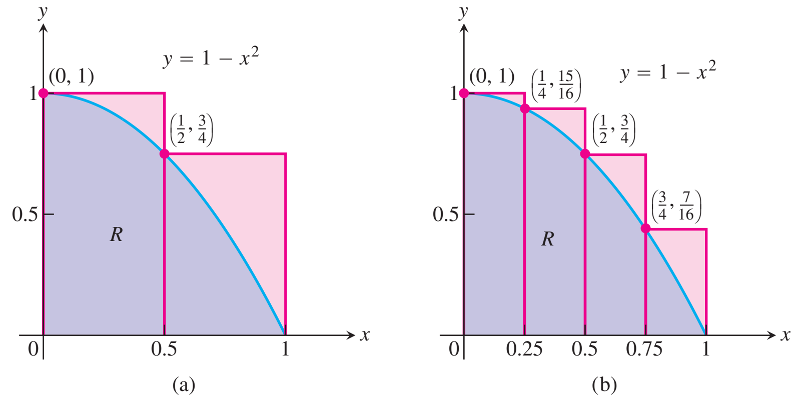


FIGURE 5.2 (a) We get an upper estimate of the area of R by using two rectangles containing R. (b) Four rectangles give a better upper estimate. Both estimates overshoot the true value for the area.

While we do not yet have a method for determining the exact area of R, we can approximate it in a simple way. Figure 5.2a shows two rectangles that together contain the region R. Each rectangle has width $1/2$ and they have heights 1 and $3/4$, moving from left to right. The height of each rectangle is the maximum value of the function f, obtained by evaluating f at the left endpoint of the subinterval of [0, 1] forming the base of the rectangle. The total area of the two rectangles approximates the area A of the region R,

$$A \approx 1 \cdot \frac{1}{2} + \frac{3}{4} \cdot \frac{1}{2} = \frac{7}{8} = 0.875.$$

This estimate is larger than the true area A, since the two rectangles contain R. We say that 0.875 is an **upper sum** because it is obtained by taking the height of each rectangle as the maximum (uppermost) value of $f(x)$ for x a point in the base interval of the rectangle. In Figure 5.2b, we improve our estimate by using four thinner rectangles, each of width $1/4$, which taken together contain the region R. These four rectangles give the approximation

$$A \approx 1 \cdot \frac{1}{4} + \frac{15}{16} \cdot \frac{1}{4} + \frac{3}{4} \cdot \frac{1}{4} + \frac{7}{16} \cdot \frac{1}{4} = \frac{25}{32} = 0.78125,$$

which is still greater than A since the four rectangles contain R.

Suppose instead we use four rectangles contained *inside* the region R to estimate the area, as in Figure 5.3a. Each rectangle has width $1/4$ as before, but the rectangles are shorter and lie entirely beneath the graph of f. The function $f(x) = 1 - x^2$ is decreasing on [0, 1], so the height of each of these rectangles is given by the value of f at the right endpoint of the subinterval forming its base. The fourth rectangle has zero height and therefore contributes no area. Summing these rectangles with heights equal to the minimum value of $f(x)$ for x a point in each base subinterval, gives a **lower sum** approximation to the area,

$$A \approx \frac{15}{16} \cdot \frac{1}{4} + \frac{3}{4} \cdot \frac{1}{4} + \frac{7}{16} \cdot \frac{1}{4} + 0 \cdot \frac{1}{4} = \frac{17}{32} = 0.53125.$$

This estimate is smaller than the area A since the rectangles all lie inside of the region R. The true value of A lies somewhere between these lower and upper sums:

$$0.53125 < A < 0.78125.$$

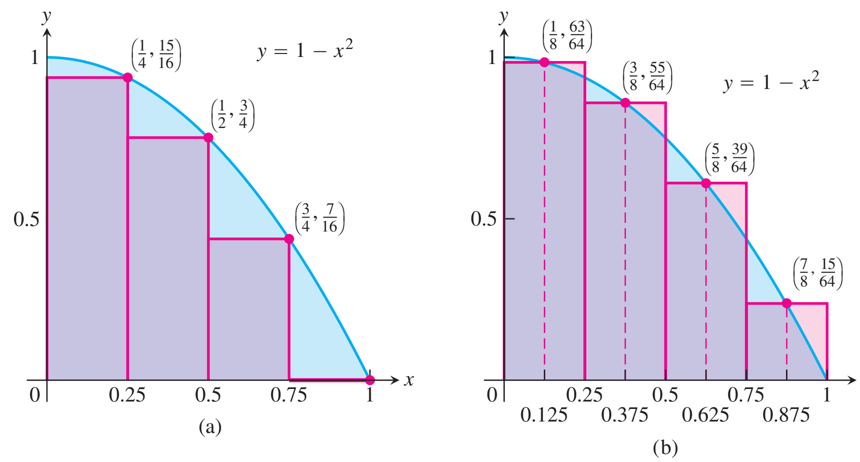


FIGURE 5.3 (a) Rectangles contained in R give an estimate for the area that undershoots the true value. (b) The midpoint rule uses rectangles whose height is the value of $y = f(x)$ at the midpoints of their bases.

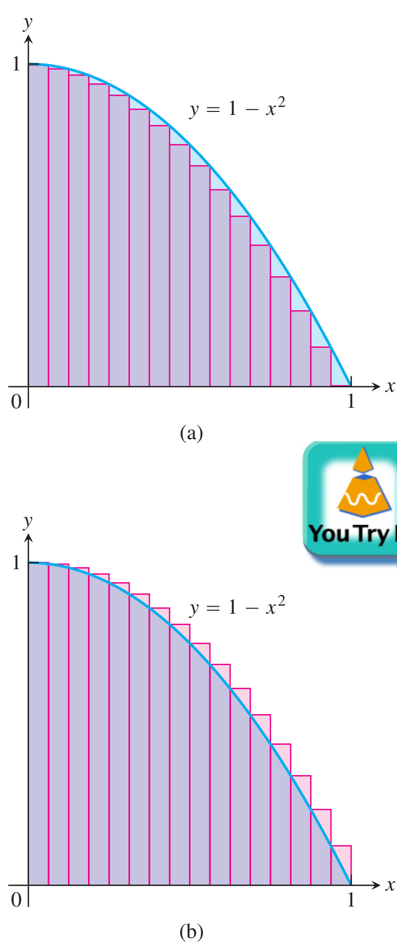


FIGURE 5.4 (a) A lower sum using 16 rectangles of equal width $\Delta x = 1/16$. (b) An upper sum using 16 rectangles.

By considering both lower and upper sum approximations we get not only estimates for the area, but also a bound on the size of the possible error in these estimates since the true value of the area lies somewhere between them. Here the error cannot be greater than the difference $0.78125 - 0.53125 = 0.25$.

Yet another estimate can be obtained by using rectangles whose heights are the values of f at the midpoints of their bases (Figure 5.3b). This method of estimation is called the **midpoint rule** for approximating the area. The midpoint rule gives an estimate that is between a lower sum and an upper sum, but it is not clear whether it overestimates or underestimates the true area. With four rectangles of width $1/4$ as before, the midpoint rule estimates the area of R to be

$$A \approx \frac{63}{64} \cdot \frac{1}{4} + \frac{55}{64} \cdot \frac{1}{4} + \frac{39}{64} \cdot \frac{1}{4} + \frac{15}{64} \cdot \frac{1}{4} = \frac{172}{64} \cdot \frac{1}{4} = 0.671875.$$

In each of our computed sums, the interval $[a, b]$ over which the function f is defined was subdivided into n subintervals of equal width (also called length) $\Delta x = (b - a)/n$, and f was evaluated at a point in each subinterval: c_1 in the first subinterval, c_2 in the second subinterval, and so on. The finite sums then all take the form

$$f(c_1)\,\Delta x + f(c_2)\,\Delta x + f(c_3)\,\Delta x + \cdots + f(c_n)\,\Delta x.$$

By taking more and more rectangles, with each rectangle thinner than before, it appears that these finite sums give better and better approximations to the true area of the region R.

Figure 5.4a shows a lower sum approximation for the area of R using 16 rectangles of equal width. The sum of their areas is 0.634765625, which appears close to the true area, but is still smaller since the rectangles lie inside R.

Figure 5.4b shows an upper sum approximation using 16 rectangles of equal width. The sum of their areas is 0.697265625, which is somewhat larger than the true area because the rectangles taken together contain R. The midpoint rule for 16 rectangles gives a total area approximation of 0.6669921875, but it is not immediately clear whether this estimate is larger or smaller than the true area.

TABLE 5.1 Finite approximations for the area of R

Number of subintervals	Lower sum	Midpoint rule	Upper sum
2	.375	.6875	.875
4	.53125	.671875	.78125
16	.634765625	.6669921875	.697265625
50	.6566	.6667	.6766
100	.66165	.666675	.67165
1000	.6661665	.66666675	.6671665

Table 5.1 shows the values of upper and lower sum approximations to the area of R using up to 1000 rectangles. In Section 5.2 we will see how to get an exact value of the areas of regions such as R by taking a limit as the base width of each rectangle goes to zero and the number of rectangles goes to infinity. With the techniques developed there, we will be able to show that the area of R is exactly $2/3$.

Distance Traveled

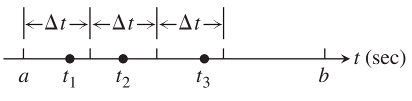
Suppose we know the velocity function $v(t)$ of a car moving down a highway, without changing direction, and want to know how far it traveled between times $t = a$ and $t = b$. If we already know an antiderivative $F(t)$ of $v(t)$ we can find the car's position function $s(t)$ by setting $s(t) = F(t) + C$. The distance traveled can then be found by calculating the change in position, $s(b) - s(a)$ (see Exercise 93, Section 4.8). If the velocity function is determined by recording a speedometer reading at various times on the car, then we have no formula from which to obtain an antiderivative function for velocity. So what do we do in this situation?

When we don't know an antiderivative for the velocity function $v(t)$, we can approximate the distance traveled in the following way. Subdivide the interval $[a, b]$ into short time intervals on each of which the velocity is considered to be fairly constant. Then approximate the distance traveled on each time subinterval with the usual distance formula

$$\text{distance} = \text{velocity} \times \text{time}$$

and add the results across $[a, b]$.

Suppose the subdivided interval looks like



with the subintervals all of equal length Δt. Pick a number t_1 in the first interval. If Δt is so small that the velocity barely changes over a short time interval of duration Δt, then the distance traveled in the first time interval is about $v(t_1)\,\Delta t$. If t_2 is a number in the second interval, the distance traveled in the second time interval is about $v(t_2)\,\Delta t$. The sum of the distances traveled over all the time intervals is

$$D \approx v(t_1)\,\Delta t + v(t_2)\,\Delta t + \cdots + v(t_n)\,\Delta t,$$

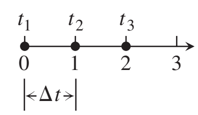
where n is the total number of subintervals.

EXAMPLE 2 Estimating the Height of a Projectile

The velocity function of a projectile fired straight into the air is $f(t) = 160 - 9.8t$ m/sec. Use the summation technique just described to estimate how far the projectile rises during the first 3 sec. How close do the sums come to the exact figure of 435.9 m?

Solution We explore the results for different numbers of intervals and different choices of evaluation points. Notice that $f(t)$ is decreasing, so choosing left endpoints gives an upper sum estimate; choosing right endpoints gives a lower sum estimate.

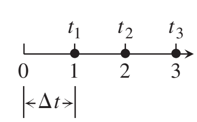
(a) *Three subintervals of length* 1, *with f evaluated at left endpoints giving an upper sum*:



With f evaluated at $t = 0, 1$, and 2, we have

$$\begin{aligned} D &\approx f(t_1)\,\Delta t + f(t_2)\,\Delta t + f(t_3)\,\Delta t \\ &= [160 - 9.8(0)](1) + [160 - 9.8(1)](1) + [160 - 9.8(2)](1) \\ &= 450.6. \end{aligned}$$

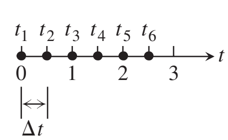
(b) *Three subintervals of length* 1, *with f evaluated at right endpoints giving a lower sum*:

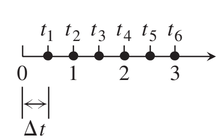


With f evaluated at $t = 1, 2$, and 3, we have

$$\begin{aligned} D &\approx f(t_1)\,\Delta t + f(t_2)\,\Delta t + f(t_3)\,\Delta t \\ &= [160 - 9.8(1)](1) + [160 - 9.8(2)](1) + [160 - 9.8(3)](1) \\ &= 421.2. \end{aligned}$$

(c) *With six subintervals of length* 1/2, *we get*





An upper sum using left endpoints: $D \approx 443.25$; a lower sum using right endpoints: $D \approx 428.55$.

These six-interval estimates are somewhat closer than the three-interval estimates. The results improve as the subintervals get shorter.

As we can see in Table 5.2, the left-endpoint upper sums approach the true value 435.9 from above, whereas the right-endpoint lower sums approach it from below. The true

TABLE 5.2 Travel-distance estimates

Number of subintervals	Length of each subinterval	Upper sum	Lower sum
3	1	450.6	421.2
6	1/2	443.25	428.55
12	1/4	439.57	432.22
24	1/8	437.74	434.06
48	1/16	436.82	434.98
96	1/32	436.36	435.44
192	1/64	436.13	435.67

value lies between these upper and lower sums. The magnitude of the error in the closest entries is 0.23, a small percentage of the true value.

$$\begin{aligned}\text{Error magnitude} &= |\text{true value} - \text{calculated value}| \\ &= |435.9 - 435.67| = 0.23.\end{aligned}$$

$$\text{Error percentage} = \frac{0.23}{435.9} \approx 0.05\%.$$

It would be reasonable to conclude from the table's last entries that the projectile rose about 436 m during its first 3 sec of flight. ■

Displacement Versus Distance Traveled

If a body with position function $s(t)$ moves along a coordinate line without changing direction, we can calculate the total distance it travels from $t = a$ to $t = b$ by summing the distance traveled over small intervals, as in Example 2. If the body changes direction one or more times during the trip, then we need to use the body's *speed* $|v(t)|$, which is the absolute value of its velocity function, $v(t)$, to find the total distance traveled. Using the velocity itself, as in Example 2, only gives an estimate to the body's **displacement**, $s(b) - s(a)$, the difference between its initial and final positions.

To see why, partition the time interval $[a, b]$ into small enough equal subintervals Δt so that the body's velocity does not change very much from time t_{k-1} to t_k. Then $v(t_k)$ gives a good approximation of the velocity throughout the interval. Accordingly, the change in the body's position coordinate during the time interval is about

$$v(t_k)\,\Delta t.$$

The change is positive if $v(t_k)$ is positive and negative if $v(t_k)$ is negative.

In either case, the distance traveled during the subinterval is about

$$|v(t_k)|\,\Delta t.$$

You Try It

The **total distance traveled** is approximately the sum

$$|v(t_1)|\,\Delta t + |v(t_2)|\,\Delta t + \cdots + |v(t_n)|\,\Delta t.$$

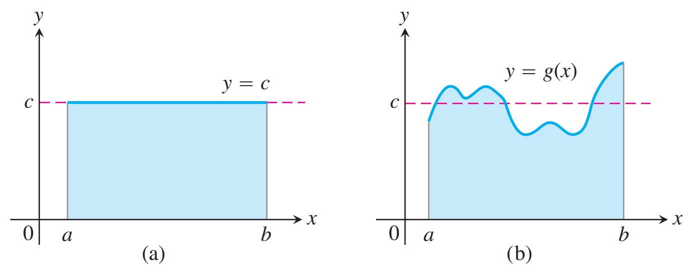


FIGURE 5.5 (a) The average value of $f(x) = c$ on $[a, b]$ is the area of the rectangle divided by $b - a$. (b) The average value of $g(x)$ on $[a, b]$ is the area beneath its graph divided by $b - a$.

Average Value of a Nonnegative Function

The average value of a collection of n numbers $x_1, x_2, \dots, x_n$ is obtained by adding them together and dividing by n. But what is the average value of a continuous function f on an interval $[a, b]$? Such a function can assume infinitely many values. For example, the temperature at a certain location in a town is a continuous function that goes up and down each day. What does it mean to say that the average temperature in the town over the course of a day is 73 degrees?

When a function is constant, this question is easy to answer. A function with constant value c on an interval $[a, b]$ has average value c. When c is positive, its graph over $[a, b]$ gives a rectangle of height c. The average value of the function can then be interpreted geometrically as the area of this rectangle divided by its width $b - a$ (Figure 5.5a).

What if we want to find the average value of a nonconstant function, such as the function g in Figure 5.5b? We can think of this graph as a snapshot of the height of some water that is sloshing around in a tank, between enclosing walls at $x = a$ and $x = b$. As the water moves, its height over each point changes, but its average height remains the same. To get the average height of the water, we let it settle down until it is level and its height is constant. The resulting height c equals the area under the graph of g divided by $b - a$. We are led to *define* the average value of a nonnegative function on an interval $[a, b]$ to be the area under its graph divided by $b - a$. For this definition to be valid, we need a precise understanding of what is meant by the area under a graph. This will be obtained in Section 5.3, but for now we look at two simple examples.

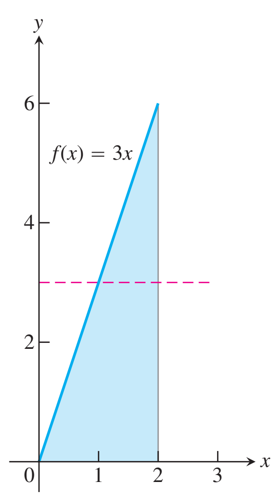


FIGURE 5.6 The average value of $f(x) = 3x$ over $[0, 2]$ is 3 (Example 3).

EXAMPLE 3 The Average Value of a Linear Function

What is the average value of the function $f(x) = 3x$ on the interval $[0, 2]$?

Solution The average equals the area under the graph divided by the width of the interval. In this case we do not need finite approximation to estimate the area of the region under the graph: a triangle of height 6 and base 2 has area 6 (Figure 5.6). The width of the interval is $b - a = 2 - 0 = 2$. The average value of the function is $6/2 = 3$. ■

EXAMPLE 4 The Average Value of $\sin x$

Estimate the average value of the function $f(x) = \sin x$ on the interval $[0, \pi]$.

Solution Looking at the graph of $\sin x$ between 0 and π in Figure 5.7, we can see that its average height is somewhere between 0 and 1. To find the average we need to

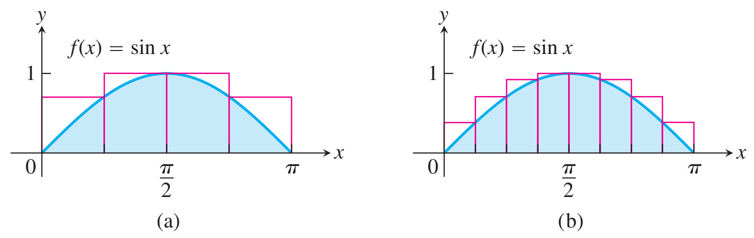


FIGURE 5.7 Approximating the area under $f(x) = \sin x$ between 0 and π to compute the average value of $\sin x$ over $[0, \pi]$, using (a) four rectangles; (b) eight rectangles (Example 4).

calculate the area A under the graph and then divide this area by the length of the interval, $\pi - 0 = \pi$.

We do not have a simple way to determine the area, so we approximate it with finite sums. To get an upper sum estimate, we add the areas of four rectangles of equal width $\pi/4$ that together contain the region beneath the graph of $y = \sin x$ and above the x-axis on $[0, \pi]$. We choose the heights of the rectangles to be the largest value of $\sin x$ on each subinterval. Over a particular subinterval, this largest value may occur at the left endpoint, the right endpoint, or somewhere between them. We evaluate $\sin x$ at this point to get the height of the rectangle for an upper sum. The sum of the rectangle areas then estimates the total area (Figure 5.7a):

$$A \approx \left(\sin\frac{\pi}{4}\right)\cdot\frac{\pi}{4} + \left(\sin\frac{\pi}{2}\right)\cdot\frac{\pi}{4} + \left(\sin\frac{\pi}{2}\right)\cdot\frac{\pi}{4} + \left(\sin\frac{3\pi}{4}\right)\cdot\frac{\pi}{4}$$
$$= \left(\frac{1}{\sqrt{2}} + 1 + 1 + \frac{1}{\sqrt{2}}\right)\cdot\frac{\pi}{4} \approx (3.42)\cdot\frac{\pi}{4} \approx 2.69.$$

To estimate the average value of $\sin x$ we divide the estimated area by π and obtain the approximation $2.69/\pi \approx 0.86$.

If we use eight rectangles of equal width $\pi/8$ all lying above the graph of $y = \sin x$ (Figure 5.7b), we get the area estimate

$$A \approx \left(\sin\frac{\pi}{8} + \sin\frac{\pi}{4} + \sin\frac{3\pi}{8} + \sin\frac{\pi}{2} + \sin\frac{\pi}{2} + \sin\frac{5\pi}{8} + \sin\frac{3\pi}{4} + \sin\frac{7\pi}{8}\right)\cdot\frac{\pi}{8}$$
$$\approx (.38 + .71 + .92 + 1 + 1 + .92 + .71 + .38)\cdot\frac{\pi}{8} = (6.02)\cdot\frac{\pi}{8} \approx 2.365.$$

Dividing this result by the length π of the interval gives a more accurate estimate of 0.753 for the average. Since we used an upper sum to approximate the area, this estimate is still greater than the actual average value of $\sin x$ over $[0, \pi]$. If we use more and more rectangles, with each rectangle getting thinner and thinner, we get closer and closer to the true average value. Using the techniques of Section 5.3, we will show that the true average value is $2/\pi \approx 0.64$.

As before, we could just as well have used rectangles lying under the graph of $y = \sin x$ and calculated a lower sum approximation, or we could have used the midpoint rule. In Section 5.3, we will see that it doesn't matter whether our approximating rectangles are chosen to give upper sums, lower sums, or a sum in between. In each case, the approximations are close to the true area if all the rectangles are sufficiently thin. ■

Summary

The area under the graph of a positive function, the distance traveled by a moving object that doesn't change direction, and the average value of a nonnegative function over an interval can all be approximated by finite sums. First we subdivide the interval into subintervals, treating the appropriate function f as if it were constant over each particular subinterval. Then we multiply the width of each subinterval by the value of f at some point within it, and add these products together. If the interval $[a, b]$ is subdivided into n subintervals of equal widths $\Delta x = (b - a)/n$, and if $f(c_k)$ is the value of f at the chosen point c_k in the kth subinterval, this process gives a finite sum of the form

$$f(c_1)\,\Delta x + f(c_2)\,\Delta x + f(c_3)\,\Delta x + \cdots + f(c_n)\,\Delta x.$$

The choices for the c_k could maximize or minimize the value of f in the kth subinterval, or give some value in between. The true value lies somewhere between the approximations given by upper sums and lower sums. The finite sum approximations we looked at improved as we took more subintervals of thinner width.