

...now is a good time to read (or
 about A Preview of Calculus (see
 2). It discusses the unifying ideas
 of calculus and helps put in perspec-
 tive where we have been and where
 we are going.

In Chapter 2 we used the tangent and velocity problems to introduce the derivative, which is the central idea in differential calculus. In much the same way, this chapter starts with the area and distance problems and uses them to formulate the idea of a definite integral, which is the basic concept of integral calculus. We will see in Chapters 6 and 8 how to use the integral to solve problems concerning volumes, lengths of curves, population predictions, cardiac output, forces on a dam, work, consumer surplus, and baseball, among many others.

There is a connection between integral calculus and differential calculus. The Fundamental Theorem of Calculus relates the integral to the derivative, and we will see in this chapter that it greatly simplifies the solution of many problems.

5.1 Areas and Distances

In this section we discover that in attempting to find the area under a curve or the distance traveled by a car, we end up with the same special type of limit.

The Area Problem

We begin by attempting to solve the *area problem*: Find the area of the region S that lies under the curve $y = f(x)$ from a to b . This means that S , illustrated in Figure 1, is bounded by the graph of a continuous function f [where $f(x) \geq 0$], the vertical lines $x = a$ and $x = b$, and the x -axis.

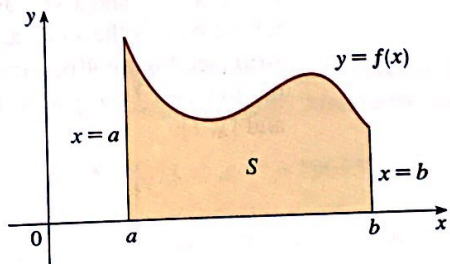


FIGURE 1

$$S = \{(x, y) \mid a \leq x \leq b, 0 \leq y \leq f(x)\}$$

In trying to solve the area problem we have to ask ourselves: What is the meaning of the word *area*? This question is easy to answer for regions with straight sides. For a rectangle, the area is defined as the product of the length and the width. The area of a triangle is half the base times the height. The area of a polygon is found by dividing it into triangles (as in Figure 2) and adding the areas of the triangles.

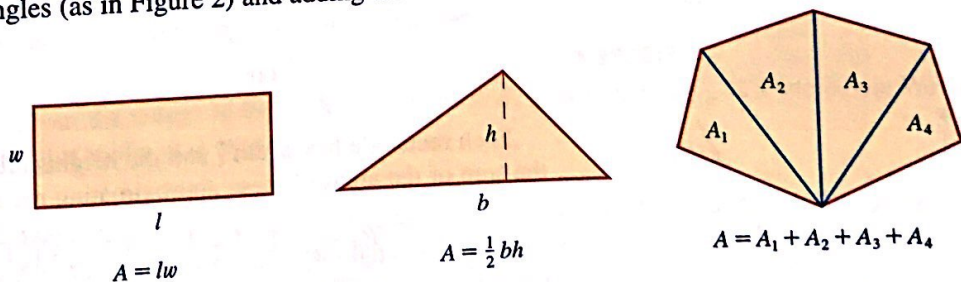


FIGURE 2

However, it is not so easy to find the area of a region with curved sides. We all have an intuitive idea of what the area of a region is. But part of the area problem is to make this intuitive idea precise by giving an exact definition of area.

Recall that in defining a tangent we first approximated the slope of the tangent line by slopes of secant lines and then we took the limit of these approximations. We pursue a similar idea for areas. We first approximate the region S by rectangles and then we take the limit of the areas of these rectangles as we increase the number of rectangles. The following example illustrates the procedure.

Try placing rectangles to estimate the area.



Resources / Module 6
/ What Is Area?
/ Estimating Area under a Parabola

EXAMPLE 1 □ Use rectangles to estimate the area under the parabola $y = x^2$ from 0 to 1 (the parabolic region S illustrated in Figure 3).

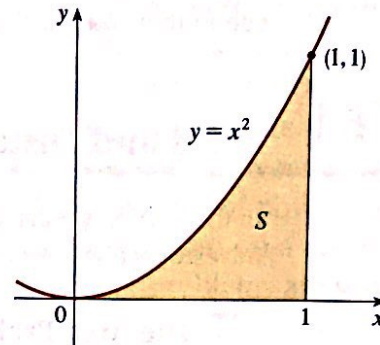


FIGURE 3

SOLUTION We first notice that the area of S must be somewhere between 0 and 1 because S is contained in a square with side length 1, but we can certainly do better than that. Suppose we divide S into four strips $S_1, S_2, S_3,$ and S_4 by drawing the vertical lines $x = \frac{1}{4}, x = \frac{1}{2},$ and $x = \frac{3}{4}$ as in Figure 4(a). We can approximate each strip by a rectangle whose base is the same as the strip and whose height is the same as the right edge of the strip [see Figure 4(b)]. In other words, the heights of these rectangles are the values of the function $f(x) = x^2$ at the right endpoints of the subintervals $[0, \frac{1}{4}], [\frac{1}{4}, \frac{1}{2}], [\frac{1}{2}, \frac{3}{4}],$ and $[\frac{3}{4}, 1]$.

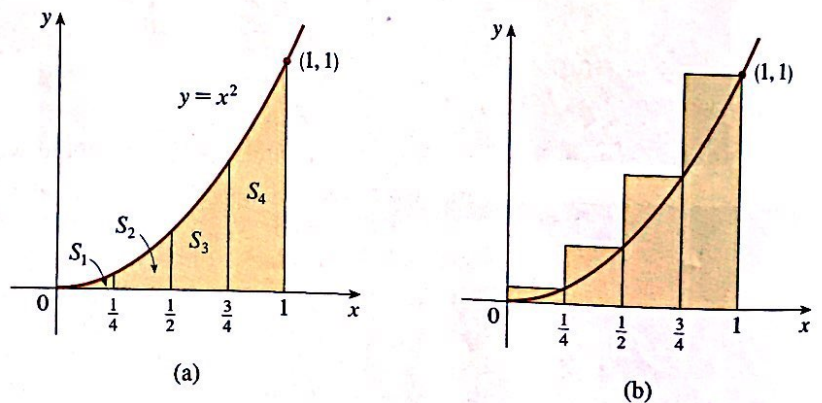


FIGURE 4

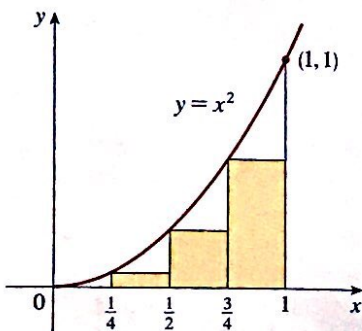


FIGURE 5

Each rectangle has width $\frac{1}{4}$ and the heights are $(\frac{1}{4})^2, (\frac{1}{2})^2, (\frac{3}{4})^2,$ and 1^2 . If we let R_4 be the sum of the areas of these approximating rectangles, we get

$$R_4 = \frac{1}{4} \cdot (\frac{1}{4})^2 + \frac{1}{4} \cdot (\frac{1}{2})^2 + \frac{1}{4} \cdot (\frac{3}{4})^2 + \frac{1}{4} \cdot 1^2 = \frac{15}{32} = 0.46875$$

From Figure 4(b) we see that the area A of S is less than $R_4,$ so

$$A < 0.46875$$

Instead of using the rectangles in Figure 4(b) we could use the smaller rectangles in Figure 5 whose heights are the values of f at the left-hand endpoints of the subintervals. (The leftmost rectangle has collapsed because its height is 0.) The sum of the areas of

these approximating rectangles is

$$L_4 = \frac{1}{4} \cdot 0^2 + \frac{1}{4} \cdot \left(\frac{1}{4}\right)^2 + \frac{1}{4} \cdot \left(\frac{1}{2}\right)^2 + \frac{1}{4} \cdot \left(\frac{3}{4}\right)^2 = \frac{7}{32} = 0.21875$$

We see that the area of S is larger than L_4 , so we have lower and upper estimates for A :

$$0.21875 < A < 0.46875$$

We can repeat this procedure with a larger number of strips. Figure 6 shows what happens when we divide the region S into eight strips of equal width.

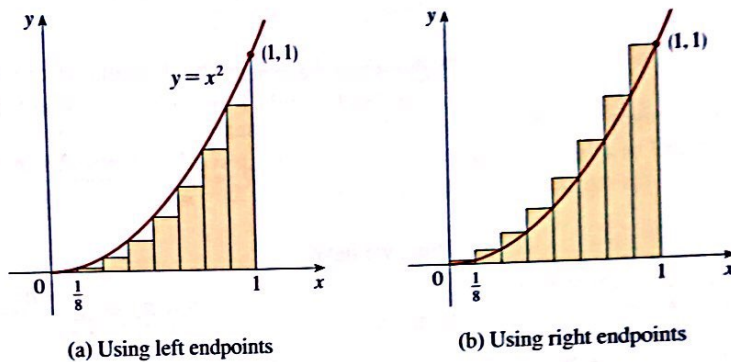


FIGURE 6

Approximating S with eight rectangles

By computing the sum of the areas of the smaller rectangles (L_8) and the sum of the areas of the larger rectangles (R_8), we obtain better lower and upper estimates for A :

$$0.2734375 < A < 0.3984375$$

So one possible answer to the question is to say that the true area of S lies somewhere between 0.2734375 and 0.3984375.

We could obtain better estimates by increasing the number of strips. The table at the left shows the results of similar calculations (with a computer) using n rectangles whose heights are found with left-hand endpoints (L_n) or right-hand endpoints (R_n). In particular, we see by using 50 strips that the area lies between 0.3234 and 0.3434. With 1000 strips we narrow it down even more: A lies between 0.3328335 and 0.3338335. A good estimate is obtained by averaging these numbers: $A \approx 0.3333335$. □

| n | L_n | R_n |
|------|-----------|-----------|
| 10 | 0.2850000 | 0.3850000 |
| 20 | 0.3087500 | 0.3587500 |
| 30 | 0.3168519 | 0.3501852 |
| 50 | 0.3234000 | 0.3434000 |
| 100 | 0.3283500 | 0.3383500 |
| 1000 | 0.3328335 | 0.3338335 |

From the values in the table it looks as if R_n is approaching $\frac{1}{3}$ as n increases. We confirm this in the next example.

EXAMPLE 2 □ For the region S in Example 1, show that the sum of the areas of the upper approximating rectangles approaches $\frac{1}{3}$, that is,

$$\lim_{n \rightarrow \infty} R_n = \frac{1}{3}$$

SOLUTION R_n is the sum of the areas of the n rectangles in Figure 7. Each rectangle has width $1/n$ and the heights are the values of the function $f(x) = x^2$ at the points $1/n, 2/n, 3/n, \dots, n/n$; that is, the heights are $(1/n)^2, (2/n)^2, (3/n)^2, \dots, (n/n)^2$.

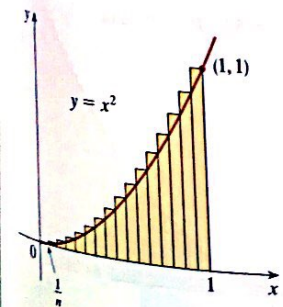


FIGURE 7

Thus

$$\begin{aligned}
 R_n &= \frac{1}{n} \left(\frac{1}{n}\right)^2 + \frac{1}{n} \left(\frac{2}{n}\right)^2 + \frac{1}{n} \left(\frac{3}{n}\right)^2 + \cdots + \frac{1}{n} \left(\frac{n}{n}\right)^2 \\
 &= \frac{1}{n} \cdot \frac{1}{n^2} (1^2 + 2^2 + 3^2 + \cdots + n^2) \\
 &= \frac{1}{n^3} (1^2 + 2^2 + 3^2 + \cdots + n^2)
 \end{aligned}$$

Here we need the formula for the sum of the squares of the first n positive integers:

$$\boxed{1} \quad 1^2 + 2^2 + 3^2 + \cdots + n^2 = \frac{n(n+1)(2n+1)}{6}$$

Perhaps you have seen this formula before. It is proved in Example 5 in Appendix E. Putting Formula 1 into our expression for R_n , we get

$$R_n = \frac{1}{n^3} \cdot \frac{n(n+1)(2n+1)}{6} = \frac{(n+1)(2n+1)}{6n^2}$$

Thus, we have

$$\begin{aligned}
 \lim_{n \rightarrow \infty} R_n &= \lim_{n \rightarrow \infty} \frac{(n+1)(2n+1)}{6n^2} \\
 &= \lim_{n \rightarrow \infty} \frac{1}{6} \left(\frac{n+1}{n}\right) \left(\frac{2n+1}{n}\right) \\
 &= \lim_{n \rightarrow \infty} \frac{1}{6} \left(1 + \frac{1}{n}\right) \left(2 + \frac{1}{n}\right) \\
 &= \frac{1}{6} \cdot 1 \cdot 2 = \frac{1}{3}
 \end{aligned}$$

□ Here we are computing the limit of the sequence $\{R_n\}$. Sequences were discussed in *A Preview of Calculus* and will be studied in detail in Chapter 11. Their limits are calculated in the same way as limits at infinity (Section 2.6). In particular, we know that

$$\lim_{n \rightarrow \infty} \frac{1}{n} = 0$$

It can be shown that the lower approximating sums also approach $\frac{1}{3}$, that is,

$$\lim_{n \rightarrow \infty} L_n = \frac{1}{3}$$

From Figures 8 and 9 it appears that, as n increases, both L_n and R_n become better and better approximations to the area of S . Therefore, we *define* the area A to be the limit of the

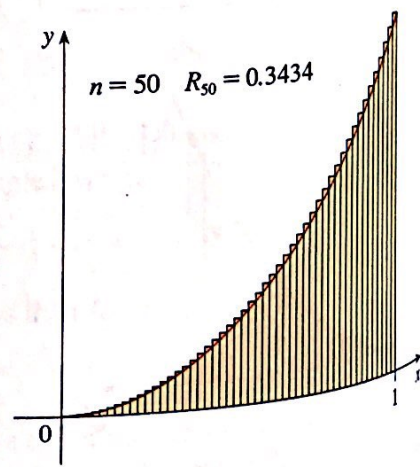
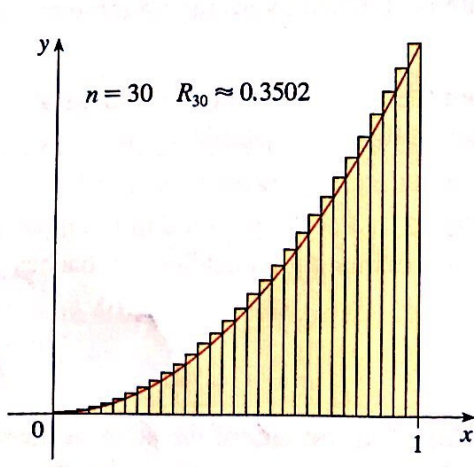
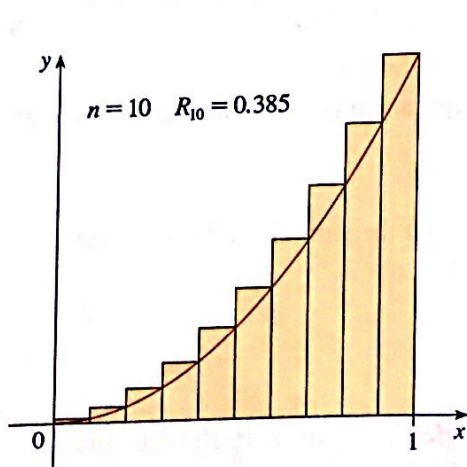
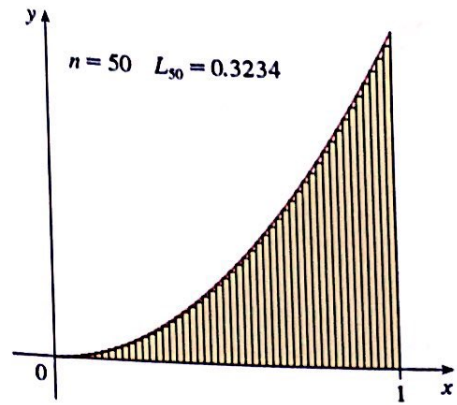
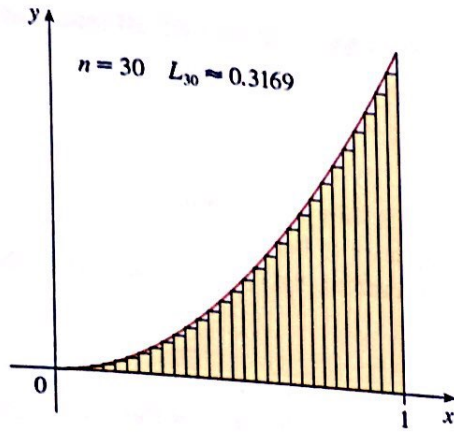
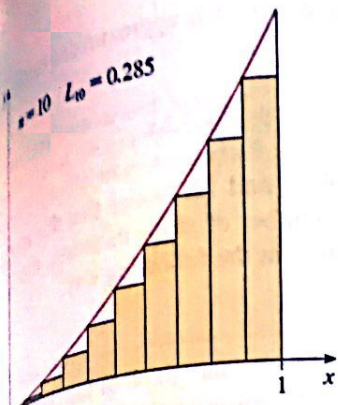


FIGURE 8



sums of the areas of the approximating rectangles, that is,

$$A = \lim_{n \rightarrow \infty} R_n = \lim_{n \rightarrow \infty} L_n = \frac{1}{3}$$

Let's apply the idea of Examples 1 and 2 to the more general region S of Figure 1. We start by subdividing S into n strips S_1, S_2, \dots, S_n of equal width as in Figure 10. The width of the interval $[a, b]$ is $b - a$, so the width of each of the n strips is

$$\Delta x = \frac{b - a}{n}$$

These strips divide the interval $[a, b]$ into n subintervals

$$[x_0, x_1], [x_1, x_2], [x_2, x_3], \dots, [x_{n-1}, x_n]$$

where $x_0 = a$ and $x_n = b$. The right-hand endpoints of the subintervals are

$$x_1 = a + \Delta x, \quad x_2 = a + 2\Delta x, \quad x_3 = a + 3\Delta x, \quad \dots$$

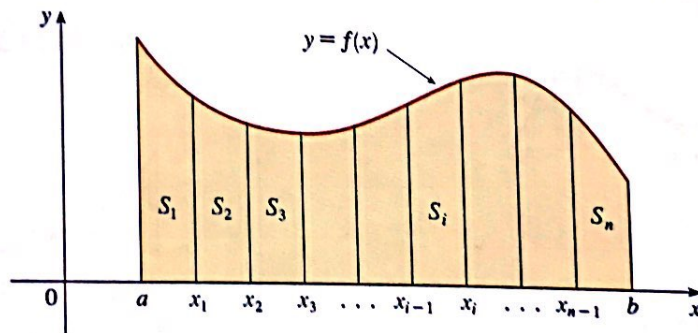


FIGURE 10

Let's approximate the i th strip S_i by a rectangle with width Δx and height $f(x_i)$, which is the value of f at the right-hand endpoint (see Figure 11). Then the area of the i th rect-

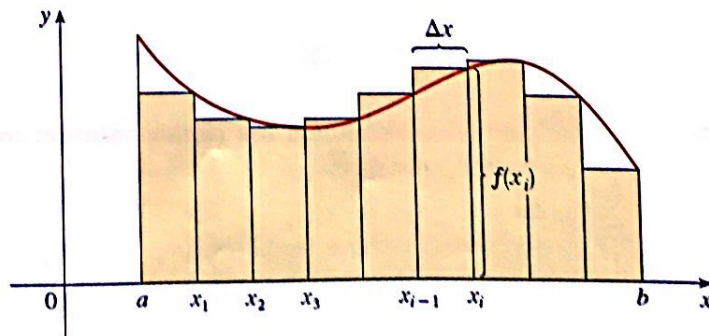
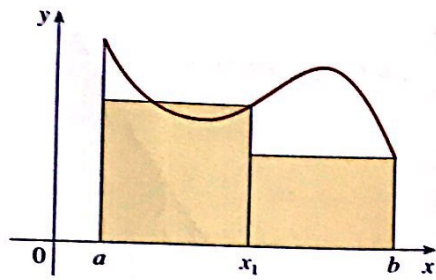
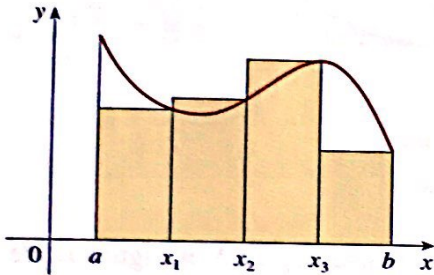


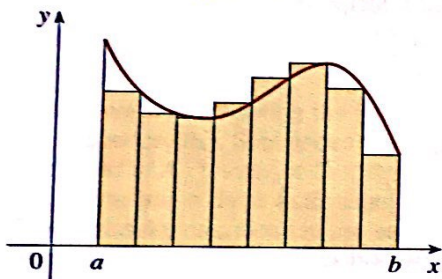
FIGURE 11



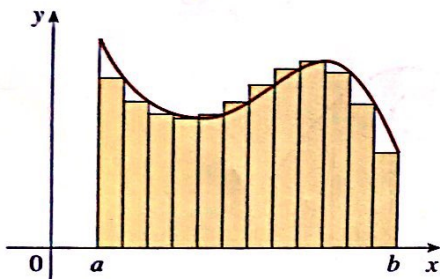
(a) $n = 2$



(b) $n = 4$



(c) $n = 8$



(d) $n = 12$

FIGURE 12

angle is $f(x_i) \Delta x$. What we think of intuitively as the area of S is approximated by the sum of the areas of these rectangles, which is

$$R_n = f(x_1) \Delta x + f(x_2) \Delta x + \cdots + f(x_n) \Delta x$$

Figure 12 shows this approximation for $n = 2, 4, 8,$ and 12 . Notice that this approximation appears to become better and better as the number of strips increases, that is, as $n \rightarrow \infty$. Therefore, we define the area A of the region S in the following way.

2 Definition The area A of the region S that lies under the graph of the continuous function f is the limit of the sum of the areas of approximating rectangles:

$$A = \lim_{n \rightarrow \infty} R_n = \lim_{n \rightarrow \infty} [f(x_1) \Delta x + f(x_2) \Delta x + \cdots + f(x_n) \Delta x]$$

It can be proved that the limit in Definition 2 always exists, since we are assuming that f is continuous. It can also be shown that we get the same value if we use left endpoints:

3
$$A = \lim_{n \rightarrow \infty} L_n = \lim_{n \rightarrow \infty} [f(x_0) \Delta x + f(x_1) \Delta x + \cdots + f(x_{n-1}) \Delta x]$$

In fact, instead of using left endpoints or right endpoints, we could take the height of the i th rectangle to be the value of f at any number x_i^* in the i th subinterval $[x_{i-1}, x_i]$. We call the numbers $x_1^*, x_2^*, \dots, x_n^*$ the **sample points**. Figure 13 shows approximating rectangles when the sample points are not chosen to be endpoints. So a more general expression for the area of S is

4
$$A = \lim_{n \rightarrow \infty} [f(x_1^*) \Delta x + f(x_2^*) \Delta x + \cdots + f(x_n^*) \Delta x]$$

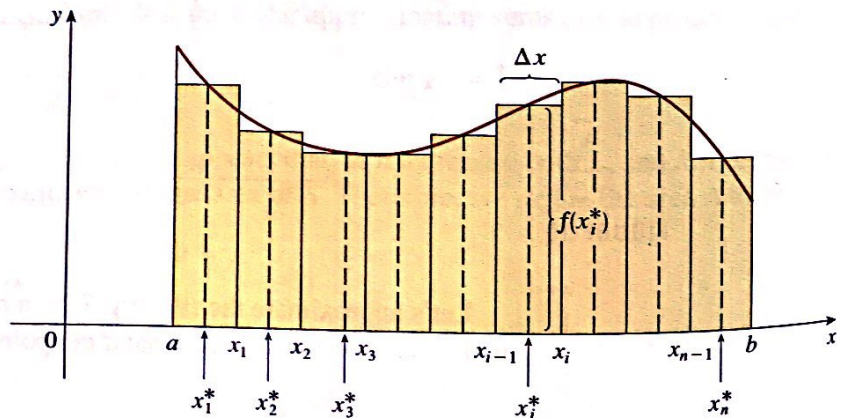


FIGURE 13

We often use **sigma notation** to write sums with many terms more compactly. For instance,

$$\sum_{i=1}^n f(x_i) \Delta x = f(x_1) \Delta x + f(x_2) \Delta x + \cdots + f(x_n) \Delta x$$

This tells us to end with $i = n$.
 This tells us to add. $\sum_{i=m}^n f(x_i) \Delta x$
 This tells us to start with $i = m$.

So the expressions for area in Equations 2, 3, and 4 can be written as follows:

$$A = \lim_{n \rightarrow \infty} \sum_{i=1}^n f(x_i) \Delta x$$

$$A = \lim_{n \rightarrow \infty} \sum_{i=1}^n f(x_{i-1}) \Delta x$$

$$A = \lim_{n \rightarrow \infty} \sum_{i=1}^n f(x_i^*) \Delta x$$

We could also rewrite Formula 1 in the following way:

$$\sum_{i=1}^n i^2 = \frac{n(n+1)(2n+1)}{6}$$

EXAMPLE 3 □ Let A be the area of the region that lies under the graph of $f(x) = e^{-x}$ between $x = 0$ and $x = 2$.

- (a) Using right endpoints, find an expression for A as a limit. Do not evaluate the limit.
 (b) Estimate the area by taking the sample points to be midpoints and using four subintervals and then ten subintervals.

SOLUTION

(a) Since $a = 0$ and $b = 2$, the width of a subinterval is

$$\Delta x = \frac{2-0}{n} = \frac{2}{n}$$

So $x_1 = 2/n$, $x_2 = 4/n$, $x_3 = 6/n$, $x_i = 2i/n$, and $x_n = 2n/n$. The sum of the areas of the approximating rectangles is

$$\begin{aligned} R_n &= f(x_1) \Delta x + f(x_2) \Delta x + \cdots + f(x_n) \Delta x \\ &= e^{-2/n} \Delta x + e^{-4/n} \Delta x + \cdots + e^{-2n/n} \Delta x \\ &= e^{-2/n} \left(\frac{2}{n} \right) + e^{-4/n} \left(\frac{2}{n} \right) + \cdots + e^{-2n/n} \left(\frac{2}{n} \right) \end{aligned}$$

According to Definition 2, the area is

$$A = \lim_{n \rightarrow \infty} R_n = \lim_{n \rightarrow \infty} \frac{2}{n} (e^{-2/n} + e^{-4/n} + e^{-6/n} + \cdots + e^{-2n/n})$$

Using sigma notation we could write

$$A = \lim_{n \rightarrow \infty} \frac{2}{n} \sum_{i=1}^n e^{-2i/n}$$

It is difficult to evaluate this limit directly by hand, but with the aid of a computer algebra system it isn't hard (see Exercise 22). In Section 5.3 we will be able to find A more easily using a different method.

(b) With $n = 4$ the subintervals of equal width $\Delta x = 0.5$ are $[0, 0.5]$, $[0.5, 1]$, $[1, 1.5]$, and $[1.5, 2]$. The midpoints of these subintervals are $x_1^* = 0.25$, $x_2^* = 0.75$, $x_3^* = 1.25$,

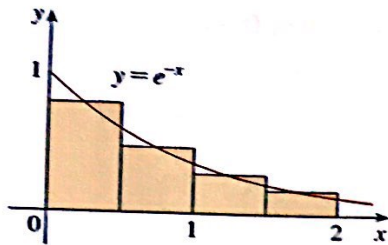


FIGURE 14

and $x_4^* = 1.75$, and the sum of the areas of the four approximating rectangles (see Figure 14) is

$$\begin{aligned} M_4 &= \sum_{i=1}^4 f(x_i^*) \Delta x \\ &= f(0.25) \Delta x + f(0.75) \Delta x + f(1.25) \Delta x + f(1.75) \Delta x \\ &= e^{-0.25}(0.5) + e^{-0.75}(0.5) + e^{-1.25}(0.5) + e^{-1.75}(0.5) \\ &= \frac{1}{2}(e^{-0.25} + e^{-0.75} + e^{-1.25} + e^{-1.75}) \approx 0.8557 \end{aligned}$$

So an estimate for the area is

$$A \approx 0.8557$$

With $n = 10$ the subintervals are $[0, 0.2], [0.2, 0.4], \dots, [1.8, 2]$ and the midpoints are $x_1^* = 0.1, x_2^* = 0.3, x_3^* = 0.5, \dots, x_{10}^* = 1.9$. Thus

$$\begin{aligned} A &\approx M_{10} = f(0.1) \Delta x + f(0.3) \Delta x + f(0.5) \Delta x + \dots + f(1.9) \Delta x \\ &= 0.2(e^{-0.1} + e^{-0.3} + e^{-0.5} + \dots + e^{-1.9}) \approx 0.8632 \end{aligned}$$

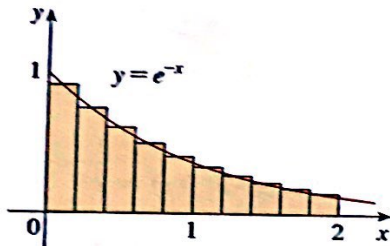


FIGURE 15

From Figure 15 it appears that this estimate is better than the estimate with $n = 4$.

The Distance Problem

Now let's consider the *distance problem*: Find the distance traveled by an object during a certain time period if the velocity of the object is known at all times. (In a sense this is the inverse problem of the velocity problem that we discussed in Section 2.1.) If the velocity remains constant, then the distance problem is easy to solve by means of the formula

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

But if the velocity varies, it is not so easy to find the distance traveled. We investigate the problem in the following example.

EXAMPLE 4 □ Suppose the odometer on our car is broken and we want to estimate the distance driven over a 30-second time interval. We take speedometer readings every five seconds and record them in the following table:

| | | | | | | | |
|-----------------|----|----|----|----|----|----|----|
| Time (s) | 0 | 5 | 10 | 15 | 20 | 25 | 30 |
| Velocity (mi/h) | 17 | 21 | 24 | 29 | 32 | 31 | 28 |

In order to have the time and the velocity in consistent units, let's convert the velocity readings to feet per second ($1 \text{ mi/h} = 5280/3600 \text{ ft/s}$):

| | | | | | | | |
|-----------------|----|----|----|----|----|----|----|
| Time (s) | 0 | 5 | 10 | 15 | 20 | 25 | 30 |
| Velocity (ft/s) | 25 | 31 | 35 | 43 | 47 | 46 | 41 |

During the first five seconds the velocity doesn't change very much, so we can estimate the distance traveled during that time by assuming that the velocity is constant. If we take the velocity during that time interval to be the initial velocity (25 ft/s), then we

obtain the approximate distance traveled during the first five seconds:

$$25 \text{ ft/s} \times 5 \text{ s} = 125 \text{ ft}$$

Similarly, during the second time interval the velocity is approximately constant and we take it to be the velocity when $t = 5$ s. So our estimate for the distance traveled from $t = 5$ s to $t = 10$ s is

$$31 \text{ ft/s} \times 5 \text{ s} = 155 \text{ ft}$$

If we add similar estimates for the other time intervals, we obtain an estimate for the total distance traveled:

$$25 \times 5 + 31 \times 5 + 35 \times 5 + 43 \times 5 + 47 \times 5 + 46 \times 5 = 1135 \text{ ft}$$

We could just as well have used the velocity at the *end* of each time period instead of the velocity at the beginning as our assumed constant velocity. Then our estimate becomes

$$31 \times 5 + 35 \times 5 + 43 \times 5 + 47 \times 5 + 46 \times 5 + 41 \times 5 = 1215 \text{ ft}$$

If we had wanted a more accurate estimate, we could have taken velocity readings every two seconds, or even every second. □

Perhaps the calculations in Example 4 remind you of the sums we used earlier to estimate areas. The similarity is explained when we sketch a graph of the velocity function of the car in Figure 16 and draw rectangles whose heights are the initial velocities for each time interval. The area of the first rectangle is $25 \times 5 = 125$, which is also our estimate for the distance traveled in the first five seconds. In fact, the area of each rectangle can be interpreted as a distance because the height represents velocity and the width represents time. The sum of the areas of the rectangles in Figure 16 is $L_6 = 1135$, which is our initial estimate for the total distance traveled.

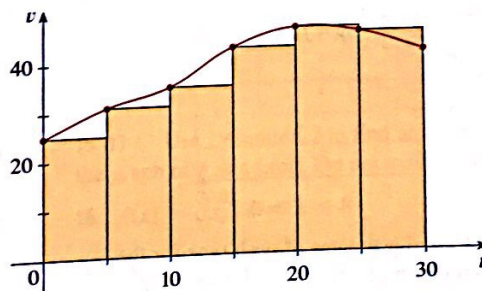


FIGURE 16

In general, suppose an object moves with velocity $v = f(t)$, where $a \leq t \leq b$ and $f(t) \geq 0$ (so the object always moves in the positive direction). We take velocity readings at times $t_0 (= a)$, t_1 , t_2 , \dots , $t_n (= b)$ so that the velocity is approximately constant on each subinterval. If these times are equally spaced, then the time between consecutive readings is $\Delta t = (b - a)/n$. During the first time interval the velocity is approximately $f(t_0)$ and so the distance traveled is approximately $f(t_0) \Delta t$. Similarly, the distance traveled during the second time interval is about $f(t_1) \Delta t$ and the total distance traveled during the time interval $[a, b]$ is approximately

$$f(t_0) \Delta t + f(t_1) \Delta t + \dots + f(t_{n-1}) \Delta t = \sum_{i=1}^n f(t_{i-1}) \Delta t$$

If we use the velocity at right-hand endpoints instead of left-hand endpoints, our estimate for the total distance becomes

$$f(t_1) \Delta t + f(t_2) \Delta t + \cdots + f(t_n) \Delta t = \sum_{i=1}^n f(t_i) \Delta t$$

The more frequently we measure the velocity, the more accurate our estimates become, so it seems plausible that the *exact* distance d traveled is the *limit* of such expressions:

$$\boxed{5} \quad d = \lim_{n \rightarrow \infty} \sum_{i=1}^n f(t_{i-1}) \Delta t = \lim_{n \rightarrow \infty} \sum_{i=1}^n f(t_i) \Delta t$$

We will see in Section 5.4 that this is indeed true.

Because Equation 5 has the same form as our expressions for area in Equations 2 and 3, it follows that the distance traveled is equal to the area under the graph of the velocity function. In Chapters 6 and 8 we will see that other quantities of interest in the natural and social sciences—such as the work done by a variable force or the cardiac output of the heart—can also be interpreted as the area under a curve. So when we compute areas in this chapter, bear in mind that they can be interpreted in a variety of practical ways.

5.1 Exercises