

Chapter 6

The Integral

There are many contexts—work, energy, area, volume, distance travelled, and profit and loss are just a few—where the quantity in which we are interested is a product of known quantities. For example, the electrical energy needed to burn three 100 watt light bulbs for Δt hours is $300 \cdot \Delta t$ watt-hours. In this example, though, the calculation becomes more complicated if lights are turned off and on during the time interval Δt . We face the same complication in any context in which one of the factors in a product varies. To describe such a product we will introduce the **integral**.

As you will see, the integral itself can be viewed as a variable quantity. By analyzing the rate at which that quantity changes, we will find that every integral can be expressed as the solution to a particular differential equation. We will thus be able to use all our tools for solving differential equations to determine integrals.

6.1 Measuring Work

Human Work

Let's measure the work done by the staff of an office that processes catalog orders. Suppose a typical worker in the office can process 10 orders an hour. Then we would expect 6 people to process 60 orders an hour; in two hours, they could process 120 orders.

Processing
catalog orders

$$10 \frac{\text{orders per hour}}{\text{person}} \times 6 \text{ persons} \times 2 \text{ hours} = 120 \text{ orders.}$$

Notice that a staff of 4 people working 3 hours could process the same number of orders:

$$10 \frac{\text{orders per hour}}{\text{person}} \times 4 \text{ persons} \times 3 \text{ hours} = 120 \text{ orders.}$$

Human work is
measured as a product

It is natural to say that 6 persons working two hours do the same amount of work as 4 persons working three hours. This suggests that we use the product

$$\text{number of workers} \times \text{elapsed time}$$

to measure **human work**. In these terms, it takes 12 “person-hours” of human work to process 120 orders.

Another name that has been used in the past for this unit of work is the “man-hour.” If the task is large, work can even be measured in “man-months” or “man-years.” The term we will use most of the time is “staff-hour.”

Productivity rate

Measuring the work in terms of person-hours or staff-hours may seem a little strange at first – after all, a typical manager of our catalog order office would be most interested in the number of orders processed; that is, the **production** of the office. Notice, however, that we can re-phrase the rate at which orders are processed as 10 orders per staff-hour. This is sometimes called the **productivity rate**. The productivity rate allows us to translate human work into production:

$$\begin{aligned} \text{production} &= \text{productivity rate} \times \text{human work} \\ 120 \text{ orders} &= 10 \frac{\text{orders}}{\text{staff-hour}} \times 12 \text{ staff-hours.} \end{aligned}$$

As this equation shows, production varies linearly with work and the productivity rate serves as multiplier (see our discussion of the multiplier on pages 31–33).

Mowing lawns

If we modify the productivity rate in a suitable way, we can use this equation for other kinds of jobs. For example, we can use it to predict how much mowing a lawn mowing crew will do. Suppose the productivity rate is .7 acres per staff-hour. Then we expect that a staff of S working for H hours can mow

$$.7 \frac{\text{acres}}{\text{staff-hour}} \times SH \text{ staff-hours} = .7 SH \text{ acres}$$

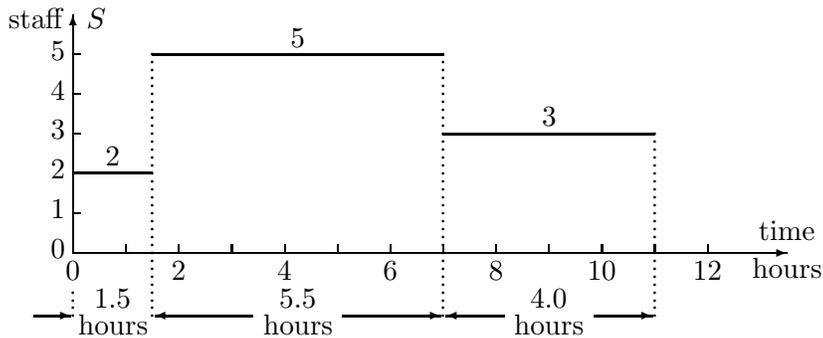
of lawn altogether.

Production is measured differently in different jobs—as orders processed, or acres mowed, or houses painted. However, in all these jobs human work is measured in the *same* way, as *staff-hours*, which gives us a common unit that can be translated from one job to another.

Staff-hours provides a common measure of work in different jobs

A staff of S working steadily for H hours does SH staff-hours of work. Suppose, though, the staffing level S is not constant, as in the graph below. Can we still find the total amount of work done?

Non-constant staffing



The basic formula works only when the staffing level is constant. But staffing *is* constant over certain time intervals. Thus, to find the total amount of work done, we should simply use the basic formula on each of those intervals, and then add up the individual contributions. These calculations are done in the following table. The total work is 42.5 staff-hours. So if the productivity rate is 10 orders per staff-hour, 425 orders can be processed.

The work done is a *sum* of products

$$2 \text{ staff} \times 1.5 \text{ hours} = 3.0 \text{ staff-hours}$$

$$5 \quad \times 5.5 \quad = 27.5$$

$$3 \quad \times 4.0 \quad = 12.0$$

$$\text{TOTAL} = 42.5 \text{ staff-hours}$$

Accumulated work

The last calculation tells us how much work got done over an entire day. What can we tell an office manager who wants to know how work is progressing *during* the day?

At the beginning of the day, only two people are working, so after the first T hours (where $0 \leq T \leq 1.5$)

$$\text{work done up to time } T = 2 \text{ staff} \times T \text{ hours} = 2T \text{ staff-hours.}$$

Even before we consider what happens after 1.5 hours, this expression calls our attention to the fact that *accumulated work is a function*—let's denote it $W(T)$. According to the formula, for the first 1.5 hours $W(T)$ is a linear function whose multiplier is

$$W' = 2 \frac{\text{staff-hours}}{\text{hour}}.$$

Work accumulates at a rate equal to the number of staff

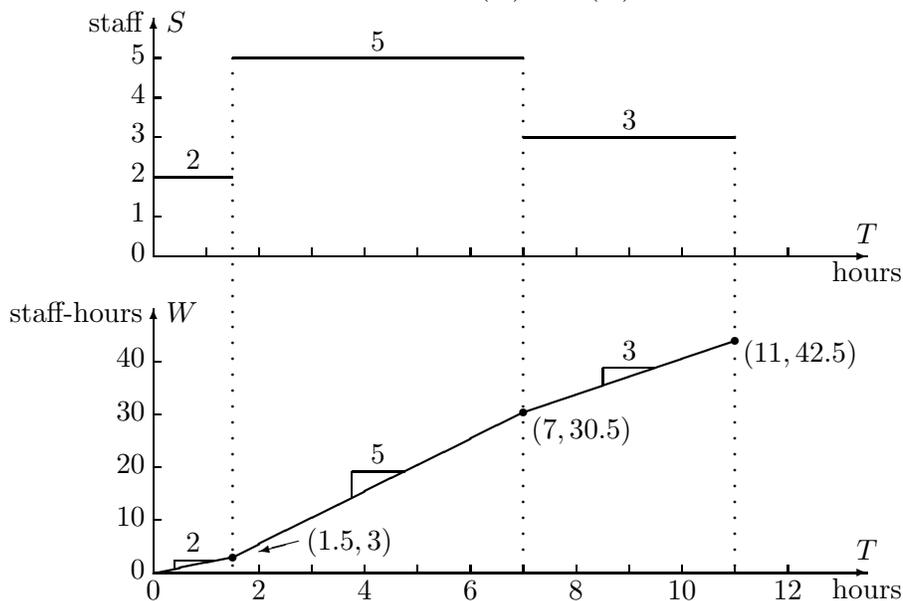
This multiplier is the **rate** at which work is being accumulated. It is also the **slope** of the graph of $W(T)$ over the interval $0 \leq T \leq 1.5$. With this insight, we can determine the rest of the graph of $W(T)$.

What must $W(T)$ look like on the next time interval $1.5 \leq T \leq 7$? Here 5 members of staff are working, so work is accumulating at the rate of 5 staff-hours per hour. Therefore, on this interval the graph of W is a straight line segment whose slope is 5 staff-hours per hour. On the third interval, the graph is another straight line segment whose slope is 3 staff-hours per hour. The complete graph of $W(T)$ is shown below.

S is the derivative of W , so ...

As the graphs show, the *slope* of the accumulated work function $W(T)$ is the *height* of the staffing function $S(T)$. In other words, S is the *derivative* of W :

$$W'(T) = S(T).$$



The accumulated work function $W(T)$

Notice that the units for W' and for S are compatible: the units for W' are “staff-hours per hour”, which we can think of as “staff”, the units for S .

We can describe the relation between S and W another way. At the moment, we have explained S in terms of W . However, since we started with S , it is really more appropriate to reverse the roles, and explain W in terms of S . Chapter 4.5 gives us the language to do this: W is an **antiderivative** of S . In other words, $y = W(T)$ is a solution to the differential equation

$$\frac{dy}{dT} = S(T).$$

As we find accumulation functions in other contexts, this relation will give us crucial information.

Before leaving this example we note some special features of S and W . The staffing function S is said to be **piecewise constant**, or a **step function**. The graphs illustrate the general fact that the derivative of a piecewise linear function (W , in this case) is piecewise constant.

... W is an antiderivative of S

The derivative of a piecewise linear function

Summary

The example of human work illustrates the key ideas we will meet, again and again, in different contexts in this chapter. Essentially, we have two functions $W(t)$ and $S(t)$ and two different ways of expressing the relation between them: On the one hand,

$$W(t) \text{ is an accumulation function for } S(t),$$

while on the other hand,

$$S(t) \text{ is the derivative of } W(t).$$

Exploring the far-reaching implications of functions connected by such a two-fold relationship will occupy the rest of this chapter.

Electrical Energy

Just as humans do work, so does electricity. A power company charges customers for the work done by the electricity it supplies, and it measures that work in a way that is strictly analogous to the way we measure human work. The work done by electricity is usually referred to as **(electrical) energy**.

For example, suppose we illuminate two light bulbs—one rated at 100 watts, the other at 60 watts. It will take the same amount of electrical energy to burn the 100-watt bulb for 3 hours as it will to burn the 60-watt bulb for 5 hours. Both will use 300 watt-hours of electricity. The power of the light bulb—measured in watts—is analogous to the number of staff working (and, in fact, workers have sometimes been called *manpower*). The time the bulb burns is analogous to the time the staff work. Finally, the product

$$\text{energy} = \text{power} \times \text{elapsed time}$$

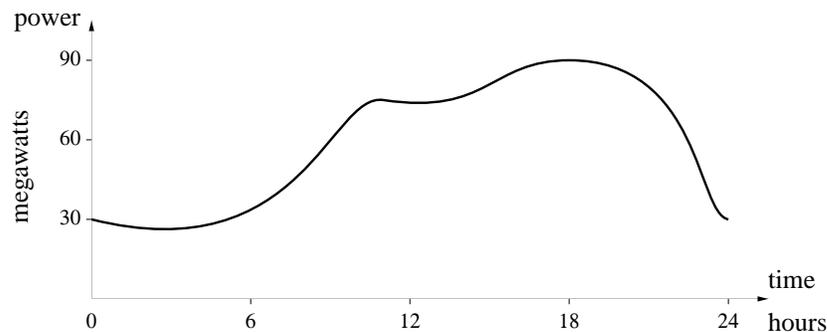
for electricity is analogous to the product

$$\text{work} = \text{number of staff} \times \text{elapsed time}$$

for human effort.

Electric *power* is measured in watts, in kilowatts (= 1,000 watts), and in megawatts (= 1,000,000 watts). Electric *energy* is measured in watt-hours, in kilowatt-hours (abbreviated ‘kwh’) and in megawatt-hours (abbreviated ‘mwh’). Since an individual electrical appliance has a power demand of about one kilowatt, kwh are suitable units to use for describing the energy consumption of a house, while mwh are more natural for a whole town.

Suppose the power demand of a town over a 24 hour period is described by the following graph:



The analogy between electrical energy and human work

Since this graph describes *power*, its vertical height over any point t on the time axis tells us the total wattage of the light bulbs, dishwashers, computers, etc. that are turned on in the town at that instant. This demand fluctuates between 30 and 90 megawatts, roughly. The problem is to determine the total amount of *energy* used in a day—how many megawatt-hours are there in this graph? Although the equation

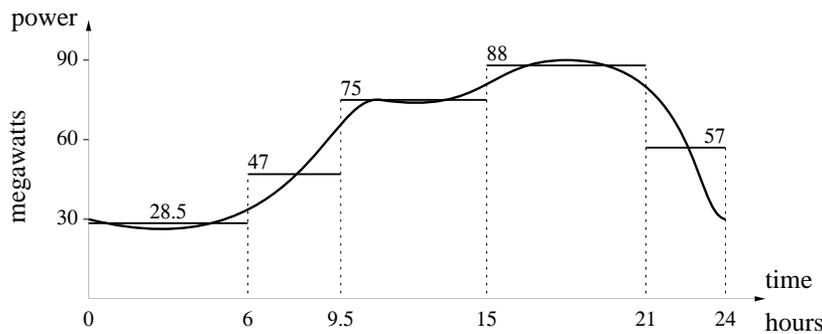
Power is analogous to staffing level

$$\text{energy} = \text{power} \times \text{elapsed time},$$

gives the basic relation between energy and power, we can't use it directly because the power demand isn't constant.

The staffing function $S(t)$ we considered earlier wasn't constant, either, but we were still able to compute staff-hours because $S(t)$ was *piecewise* constant. This suggests that we should replace the power graph by a piecewise constant graph that **approximates** it. Here is one such approximation:

A piecewise constant approximation



As you can see, the step function has five steps, so our approximation to the total energy consumption of the town will be a sum of five individual products:

$$\text{energy} \approx 28.5 \times 6 + 47 \times 3.5 + \cdots + 57 \times 3 = 1447 \text{ mwh.}$$

This value is only an *estimate*, though. How can we get a better estimate? The answer is clear: start with a step function that approximates the power graph *more closely*. In principle, we can get as good an approximation as we might desire this way. We are limited only by the precision of the power graph itself. As our approximation to the power graph improves, so does the accuracy of the calculation that estimates energy consumption.

Better estimates

In summary, we determine the energy consumption of the town by a sequence of successive approximations. The steps in the sequence are listed in the box below.

1. **Approximate** the power demand by a step function.
2. **Estimate** energy consumption from this approximation.
3. **Improve** the energy estimate by choosing a new step function that follows power demand more closely.

Accumulated energy consumption

Energy accumulation

Energy is being consumed steadily over the entire day; can we determine how much energy has been used through the first T hours of the day? We'll denote this quantity $E(T)$ and call it the **energy accumulation function**. For example, we already have the estimate $E(24) = 1447$ mwh; can we estimate $E(3)$ or $E(17.6)$?

Once again, the earlier example of human effort can guide us. We saw that work accumulates at a rate equal to the number of staff present:

$$W'(T) = S(T).$$

Since $S(T)$ was piecewise constant, this rate equation allowed us to determine $W(T)$ as a piecewise linear function.

We claim that there is an analogous relation between accumulated energy consumption and power demand—namely

$$E'(T) = p(T).$$

Unlike $S(T)$, the function $p(T)$ is not piecewise constant. Therefore, the argument we used to show that $W'(T) = S(T)$ will not work here. We need another argument.

Estimating $E'(T)$

To explain why the differential equation $E'(T) = p(T)$ should be true, we will start by analyzing

$$E'(T) \approx \frac{\Delta E}{\Delta T} = \frac{E(T + \Delta T) - E(T)}{\Delta T}.$$

Assume we have made ΔT so small that, to the level of precision we require, the approximation $\Delta E/\Delta T$ agrees with $E'(T)$. The numerator ΔE is, by

definition, the total energy used up to time $T + \Delta T$, minus the total energy used up to time T . This is just the energy used during the time interval ΔT that runs from time T to time $T + \Delta T$:

$$\Delta E = \text{energy used between times } T \text{ and } T + \Delta T.$$

Since the elapsed time ΔT is small, the power demand should be nearly constant, so we can get a good estimate for energy consumption from the basic equation

$$\text{energy used} = \text{power} \times \text{elapsed time}.$$

In particular, if we represent the power by $p(T)$, which is the power demand at the beginning of the time period from T to $T + \Delta T$, then we have

$$\Delta E \approx p(T) \cdot \Delta T.$$

Using this value in our approximation for the derivative $E'(T)$, we get

$$E'(T) \approx \frac{\Delta E}{\Delta T} \approx \frac{p(T) \cdot \Delta T}{\Delta T} = p(T).$$

That is, $E'(T) \approx p(T)$, and the approximation becomes more and more exact as the time interval ΔT shrinks to 0. Thus,

$$E'(T) = \lim_{\Delta T \rightarrow 0} \frac{\Delta E}{\Delta T} = p(T).$$

Here is another way to arrive at the same conclusion. Our starting point is the basic formula

$$\Delta E \approx p(T) \cdot \Delta T,$$

which holds over a small time interval ΔT . This formula tells us how E responds to small changes in T . But that is exactly what the **microscope equation** tells us:

$$\Delta E \approx E'(T) \cdot \Delta T.$$

Since these equations give the same information, their multipliers must be the same:

$$p(T) = E'(T).$$

In words, the differential equation $E' = p$ says that *power is the rate at which energy is consumed*. In purely mathematical terms:

A second way
to see that $E' = p$

The energy accumulation function $y = E(t)$
is a solution to the differential equation $dy/dt = p(t)$.

In fact, $y = E(t)$ is *the* solution to the **initial value problem**

$$\frac{dy}{dt} = p(t) \quad y(0) = 0.$$

We can use all the methods described in chapter 4.5 to solve this problem.

The relation we have explored between power and energy can be found in an analogous form in many other contexts, as we will see in the next two sections. In section 4 we will turn back to accumulation functions and investigate them as solutions to differential equations. Then, in chapter 11, we will look at some special methods for solving the particular differential equations that arise in accumulation problems.

Exercises

Human work

1. House-painting is a job that can be done by several people working simultaneously, so we can measure the amount of work done in “staff-hours.” Consider a house-painting business run by some students. Because of class schedules, different numbers of students will be painting at different times of the day. Let $S(T)$ be the number of staff present at time T , measured in hours from 8 am, and suppose that during an 8-hour work day, we have

$$S(T) = \begin{cases} 3, & 0 \leq T < 2, \\ 2, & 2 \leq T < 4.5, \\ 4, & 4.5 \leq T \leq 8. \end{cases}$$

- a) Draw the graph of the step function defined here, and compute the total number of staff hours.
- b) Draw the graph that shows how staff-hours *accumulate* on this job. This is the graph of the **accumulated work** function $W(T)$. (Compare the graphs of staff and staff-hours on page 340.)
- c) Determine the derivative $W'(T)$. Is $W'(T) = S(T)$?

2. Suppose that there is a house-painting job to be done, and by past experience the students know that four of them could finish it in 6 hours. But for the first 3.5 hours, only two students can show up, and after that, five will be available.

- a) How long will the whole job take? [Answer: 6.9 hours.]
- b) Draw a graph of the staffing function for this problem. Mark on the graph the time that the job is finished.
- c) Draw the graph of the accumulated work function $W(T)$.
- d) Determine the derivative $W'(T)$. Is $W'(T) = S(T)$?

Average staffing. Suppose a job can be done in three hours when 6 people work the first hour and 9 work during the last two hours. Then the job takes 24 staff-hours of work, and the **average staffing** is

$$\text{average staffing} = \frac{24 \text{ staff-hours}}{3 \text{ hours}} = 8 \text{ staff.}$$

This means that a *constant* staffing level of 8 persons can accomplish the job in the same time that the given variable staffing level did. Note that the average staffing level (8 persons) is *not* the average of the two numbers 9 and 6!

3. What is the average staffing of the jobs considered in exercises 1 and 2, above?

4. a) Draw the graph that shows how work would accumulate in the job described in exercise 1 if the work-force was kept at the *average* staffing level instead of the varying level described in the exercise. Compare this graph to the graph you drew in exercise 1 b.

b) What is the derivative $W'(T)$ of the work accumulation function whose graph you drew in part (a)?

5. What is the average staffing for the job described by the graph on page 339?

Electrical energy

6. On Monday evening, a 1500 watt space heater is left on from 7 until 11 pm. How many kilowatt-hours of electricity does it consume?

7. a) That same heater also has settings for 500 and 1000 watts. Suppose that on Tuesday we put it on the 1000 watt setting from 6 to 8 pm, then switch to 1500 watts from 8 till 11 pm, and then on the 500 watt setting through the night until 8 am, Wednesday. How much energy is consumed (in kwh)?
- b) Sketch the graphs of power demand $p(T)$ and accumulated energy consumption $E(T)$ for the space heater from Tuesday evening to Wednesday morning. Determine whether $E'(T) = p(T)$ in this case.
- c) The **average power demand** of the space heater is defined by:

$$\text{average power demand} = \frac{\text{energy consumption}}{\text{elapsed time}}.$$

If energy consumption is measured in kilowatt-hours, and time in hours, then we can measure average power demand in kilowatts—the same as power itself. (Notice the similarity with average staffing.) What is the average power demand from Tuesday evening to Wednesday morning? If the heater could be set at this average power level, how would the energy consumption compare to the actual energy consumption you determined in part (a)?

8. The graphs on pages 342 and 343 describe the power demand of a town over a 24-hour period. Give an estimate of the average power demand of the town during that period. Explain what you did to produce your estimate. [Answer: 60.29 megawatts is one estimate.]

Work as force \times distance

The effort it takes to move an object is also called work. Since it takes *twice* as much effort to move the object twice as far, or to move another object that is twice as heavy, we can see that the work done in moving an object is proportional to both the force applied and to the distance moved. The simplest way to express this fact is to define

$$\text{work} = \text{force} \times \text{distance}.$$

For example, to lift a weight of 20 pounds straight up it takes 20 pounds of force. If the vertical distance is 3 feet then

$$20 \text{ pounds} \times 3 \text{ feet} = 60 \text{ foot-pounds}$$

of work is done. Thus, once again the quantity we are interested in has the form of a product. The *foot-pound* is one of the standard units for measuring work.

9. Suppose a tractor pulls a loaded wagon over a road whose steepness varies. If the first 150 feet of road are relatively level and the tractor has to exert only 200 pounds of force while the next 400 feet are inclined and the tractor has to exert 550 pounds of force, how much work does the tractor do altogether?

10. A motor on a large ship is lifting a 2000 pound anchor that is already out of the water at the end of a 30 foot chain. The chain weighs 40 pounds per foot. As the motor lifts the anchor, the part of the chain that is hanging gets shorter and shorter, thereby reducing the weight the motor must lift.

a) What is the combined weight of anchor and hanging chain when the anchor has been lifted x feet above its initial position?

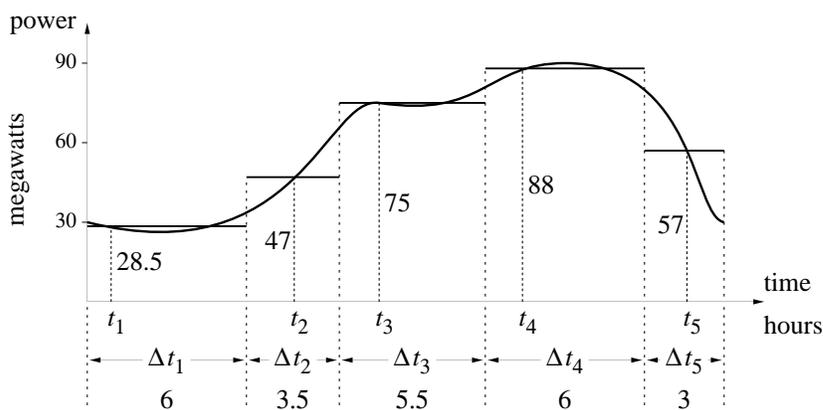
b) Divide the 30-foot distance that the anchor must move into 3 equal intervals of 10 feet each. Estimate how much work the motor does lifting the anchor and chain over each 10-foot interval by multiplying the combined weight at the *bottom* of the interval by the 10-foot height. What is your estimate for the total work done by the motor in raising the anchor and chain 30 feet?

c) Repeat all the steps of part (b), but this time use 30 equal intervals of 1 foot each. Is your new estimate of the work done larger or smaller than your estimate in part (b)? Which estimate is likely to be more accurate? On what do you base your judgment?

d) If you ignore the weight of the chain entirely, what is your estimate of the work done? How much *extra* work do you therefore estimate the motor must do to raise the heavy chain along with the anchor?

6.2 Riemann Sums

In the last section we estimated energy consumption in a town by replacing the power function $p(t)$ by a step function. Let's pause to describe that process in somewhat more general terms that we can adapt to other contexts. The power graph, the approximating step function, and the energy estimate are shown below.



$$\text{energy} \approx 28.5 \times 6 + 47 \times 3.5 + \cdots + 57 \times 3 = 1447 \text{ mwh.}$$

The height of the first step is 28.5 megawatts. This is the *actual* power level at the time t_1 indicated on the graph. That is, $p(t_1) = 28.5$ megawatts. We found a power level of 28.5 megawatts by **sampling** the power function at the time t_1 . The height of the first step could have been different if we had sampled the power function at a different time. In general, if we sample the power function $p(t)$ at the time t_1 in the interval Δt_1 , then we would estimate the energy used during that time to be

$$\text{energy} \approx p(t_1) \cdot \Delta t_1 \text{ mwh.}$$

Notice that t_1 is not in the middle, or at either end, of the first interval. It is simply a time when the power demand is representative of what's happening over the entire interval. Furthermore, t_1 is not even unique; there is another sampling time (near $t = 5$ hours) when the power level is again 28.5 megawatts.

We can describe what happens in the other time intervals the same way. If we sample the k -th interval at the point t_k , then the height of the k -th power step will be $p(t_k)$ and our estimate for the energy used during that

Sampling the
power function

time will be

$$\text{energy} \approx p(t_k) \cdot \Delta t_k \text{ mwh.}$$

We now have a general way to construct an approximation for the power function and an estimate for the energy consumed over a 24-hour period. It involves these steps.

A procedure for approximating power level and energy use

1. Choose any number n of subintervals, and let them have arbitrary positive widths $\Delta t_1, \Delta t_2, \dots, \Delta t_n$, subject only to the condition

$$\Delta t_1 + \dots + \Delta t_n = 24 \text{ hours.}$$

2. Sample the k -th subinterval at any point t_k , and let $p(t_k)$ represent the power level over this subinterval.
3. Estimate the energy used over the 24 hours by the sum

$$\text{energy} \approx p(t_1) \cdot \Delta t_1 + p(t_2) \cdot \Delta t_2 + \dots + p(t_n) \cdot \Delta t_n \text{ mwh.}$$

The expression on the right is called a **Riemann sum** for the power function $p(t)$ on the interval $0 \leq t \leq 24$ hours.

The work of Bernhard Riemann (1826–1866) has had a profound influence on contemporary mathematicians and physicists. His revolutionary ideas about the geometry of space, for example, are the basis for Einstein's theory of general relativity.

The enormous range of choices in this process means there are innumerable ways to construct a Riemann sum for $p(t)$. However, we are not really interested in *arbitrary* Riemann sums. On the contrary, we want to build Riemann sums that will give us good estimates for energy consumption. Therefore, we will choose each subinterval Δt_k so small that the power demand over that subinterval differs only very little from the sampled value $p(t_k)$. A Riemann sum constructed with *these* choices will then differ only very little from the total energy used during the 24-hour time interval.

Choices that lead to good estimates

Essentially, we use a Riemann sum to resolve a dilemma. We know the basic formula

$$\text{energy} = \text{power} \times \text{time}$$

works when power is constant, but in general power *isn't* constant—that's the dilemma. We resolve the dilemma by using instead a *sum* of terms of the form *power* \times *time*. With this sum we get an estimate for the energy.

The dilemma

In this section we will explore some other problems that present the same dilemma. In each case we will start with a basic formula that involves a product of two constant factors, and we will need to adapt the formula to the situation where one of the factors varies. The solution will be to construct a Riemann sum of such products, producing an estimate for the quantity we were after in the first place. As we work through each of these problems, you should pause to compare it to the problem of energy consumption.

Calculating Distance Travelled

Estimating velocity
and distance

It is easy to tell how far a car has travelled by reading its odometer. The problem is more complicated for a ship, particularly a sailing ship in the days before electronic navigation was common. The crew always had instruments that could measure—or at least estimate—the velocity of the ship at any time. Then, during any time interval in which the ship’s velocity is constant, the distance travelled is given by the familiar formula

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

If the velocity is *not* constant, then this formula does not work. The remedy is to break up the long time period into several short ones. Suppose their lengths are $\Delta t_1, \Delta t_2, \dots, \Delta t_n$. By assumption, the velocity is a function of time t ; let’s denote it $v(t)$. At some time t_k during each time period Δt_k measure the velocity: $v_k = v(t_k)$. Then the Riemann sum

Sampling the
velocity function

$$v(t_1) \cdot \Delta t_1 + v(t_2) \cdot \Delta t_2 + \dots + v(t_n) \cdot \Delta t_n$$

is an estimate for the total distance travelled.

For example, suppose the velocity is measured five times during a 15 hour trip—once every three hours—as shown in the table below. Then the basic formula

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

gives us an estimate for the distance travelled during each three-hour period, and the sum of these distances is an estimate of the total distance travelled during the fifteen hours. These calculations appear in the right-hand column of the table. (Note that the first measurement is used to calculate the distance travelled between hours 0 and 3, while the last measurement, taken 12 hours after the start, is used to calculate the distance travelled between hours 12 and 15.)

sampling time (hours)	elapsed time (hours)	velocity (miles/hour)	distance travelled (miles)
0	3	1.4	$3 \times 1.4 = 4.20$
3	3	5.25	$3 \times 5.25 = 15.75$
6	3	4.3	$3 \times 4.3 = 12.90$
9	3	4.6	$3 \times 4.6 = 13.80$
12	3	5.0	$3 \times 5.0 = 15.00$
			61.65

Thus we estimate the ship has travelled 61.65 miles during the fifteen hours. The number 61.65, obtained by adding the numbers in the right-most column, is a Riemann sum for the velocity function.

The estimated distance is a Riemann sum for the velocity function

Consider the specific choices that we made to construct this Riemann sum:

$$\Delta t_1 = \Delta t_2 = \Delta t_3 = \Delta t_4 = \Delta t_5 = 3$$

$$t_1 = 0, \quad t_2 = 3, \quad t_3 = 6, \quad t_4 = 9, \quad t_5 = 12.$$

These choices differ from the choices we made in the energy example in two notable ways. First, all the subintervals here are the same size. This is because it is natural to take velocity readings at regular time intervals. By contrast, in the energy example the subintervals were of different widths. Those widths were chosen in order to make a piecewise constant function that followed the power demand graph closely. Second, all the sampling times lie at the beginning of the subintervals in which they appear. Again, this is natural and convenient for velocity measurements. In the energy example, the sampling times were chosen with an eye to the power graph. Even though we can make arbitrary choices in constructing a Riemann sum, we will do it systematically whenever possible. This means choosing subintervals of equal size and sampling points at the “same” place within each interval.

Intervals and sampling times are chosen in a systematic way

Let’s turn back to our estimate for the total distance. Since the velocity of the ship could have fluctuated significantly during each of the three-hour periods we used, our estimate is rather rough. To improve the estimate we could measure the velocity more frequently—for example, every 15 minutes. If we did, the Riemann sum would have 60 terms (four distances per hour for 15 hours). The individual terms in the sum would all be much smaller, though, because they would be estimates for the distance travelled in 15

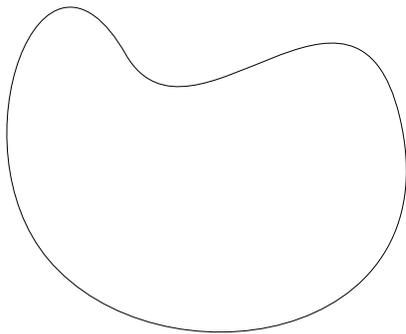
Improving the distance estimate

minutes instead of in 3 hours. For instance, the first of the 60 terms would be

$$1.4 \frac{\text{miles}}{\text{hour}} \times .25 \text{ hours} = .35 \text{ miles.}$$

Of course it may not make practical sense to do such a precise calculation. Other factors, such as water currents or the inaccuracy of the velocity measurements themselves, may keep us from getting a good estimate for the distance. Essentially, the Riemann sum is only a *model* for the distance covered by a ship.

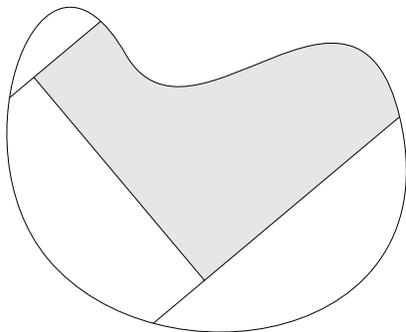
Calculating Areas



The area of a rectangle is just the product of its length and its width. How can we measure the area of a region that has an irregular boundary, like the one at the left? We would like to use the basic formula

$$\text{area} = \text{length} \times \text{width}.$$

However, since the region doesn't have straight sides, there is nothing we can call a "length" or a "width" to work with.



We can begin to deal with this problem by breaking up the region into smaller regions that *do* have straight sides—with, at most, only one curved side. This can be done many different ways. The lower figure shows one possibility. The sum of the areas of all the little regions will be the area we are looking for. Although we haven't yet solved the original problem, we have at least **reduced** it to another problem that looks simpler and may be easier to solve. Let's now work on the reduced problem for the shaded region.

Here is the shaded region, turned so that it sits flat on one of its straight sides. We would like to calculate its area using the formula

$$\text{width} \times \text{height},$$

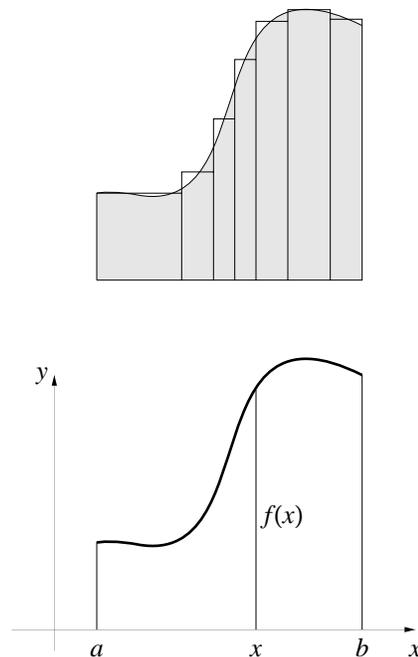
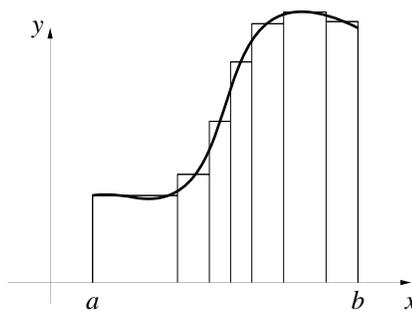
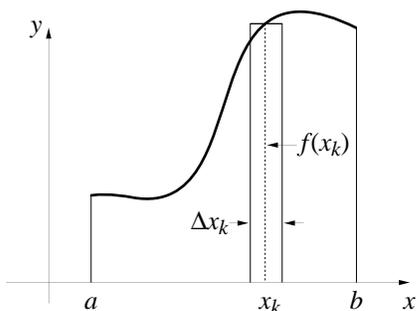
but this formula applies only to rectangles. We can, however, approximate the region by a collection of rectangles, as shown at the right. The formula *does* apply to the individual rectangles and the sum of their areas will approximate the area of the whole region.

To get the area of a rectangle, we must measure its width and height. Their heights vary with the height of the curved top of the shaded region. To describe that height in a systematic way, we have placed the shaded region in a coordinate plane so that it sits on the x -axis. The other two straight sides lie on the vertical lines $x = a$ and $x = b$.

The curved side defines the graph of a function $y = f(x)$. Therefore, at each point x , the vertical height from the axis to the curve is $f(x)$. By introducing a coordinate plane we gain access to mathematical tools—such as the language of functions—to describe the various areas.

The k -th rectangle has been singled out on the left, below. We let Δx_k denote the width of its base. By **sampling** the function f at a properly chosen point x_k in the base, we get the height $f(x_k)$ of the rectangle. Its area is therefore $f(x_k) \cdot \Delta x_k$. If we do the same thing for all n rectangles shown on the right, we can write their total area as

$$f(x_1) \cdot \Delta x_1 + f(x_2) \cdot \Delta x_2 + \cdots + f(x_n) \cdot \Delta x_n.$$



Calculating the areas
of the rectangles

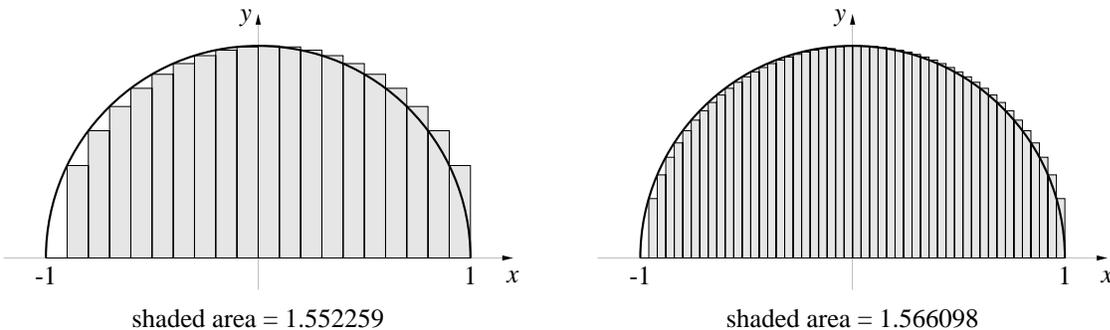
The area estimate is a Riemann sum

Notice that our estimate for the area has the form of a Riemann sum for the height function $f(x)$ over the interval $a \leq x \leq b$. To get a better estimate, we should use narrower rectangles, and more of them. In other words, we should construct another Riemann sum in which the number of terms, n , is larger and the width Δx_k of every subinterval is smaller. Putting it yet another way, we should *sample* the height more often.

Consider what happens if we apply this procedure to a region whose area we know already. The semicircle of radius $r = 1$ has an area of $\pi r^2/2 = \pi/2 = 1.5707963\dots$. The semicircle is the graph of the function

$$f(x) = \sqrt{1 - x^2},$$

which lies over the interval $-1 \leq x \leq 1$. To get the figure on the left, we sampled the height $f(x)$ at 20 evenly spaced points, starting with $x = -1$. In the better approximation on the right, we increased the number of sample points to 50. The values of the shaded areas were calculated with the program RIEMANN, which we will develop later in this section. Note that with 50 rectangles the Riemann sum is within .005 of $\pi/2$, the exact value of the area.



Calculating Lengths

It is to be expected that products—and ultimately, Riemann sums—will be involved in calculating areas. It is more surprising to find that we can use them to calculate *lengths*, too. In fact, when we are working in a coordinate plane, using a product to describe the length of a straight line is even quite natural.

To see how this can happen, consider a line segment in the x, y -plane that has a known slope m . If we also know the horizontal separation between the two ends, we can find the length of the segment. Suppose the horizontal separation is Δx and the vertical separation Δy . Then the length of the segment is

$$\sqrt{\Delta x^2 + \Delta y^2}$$

by the Pythagorean theorem (see page 90). Since $\Delta y = m \cdot \Delta x$, we can rewrite this as

$$\sqrt{\Delta x^2 + (m \cdot \Delta x)^2} = \Delta x \cdot \sqrt{1 + m^2}.$$

In other words, if a line has slope m and it is Δx units wide, then its length is the product

$$\sqrt{1 + m^2} \cdot \Delta x.$$

Suppose the line is *curved*, instead of straight. Can we describe its length the same way? We'll assume that the curve is the graph $y = g(x)$. The complication is that the slope $m = g'(x)$ now varies with x .

If $g'(x)$ doesn't vary too much over an interval of length Δx , then the curve is nearly straight. Pick a point x_* in that interval and sample the slope $g'(x_*)$ there; we expect the length of the curve to be approximately

$$\sqrt{1 + (g'(x_*))^2} \cdot \Delta x.$$

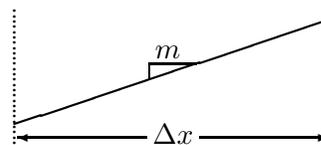
As the figure shows, this is the exact length of the straight line segment that lies over the same interval Δx and is tangent to the curve at the point $x = x_*$.

If the slope $g'(x)$ varies appreciably over the interval, we should subdivide the interval into small pieces $\Delta x_1, \Delta x_2, \dots, \Delta x_n$, over which the curve is nearly straight. Then, if we sample the slope at the point x_k in the k -th subinterval, the sum

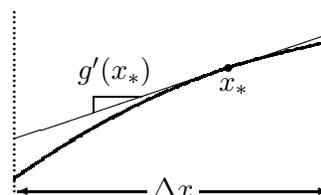
$$\sqrt{1 + (g'(x_1))^2} \cdot \Delta x_1 + \dots + \sqrt{1 + (g'(x_n))^2} \cdot \Delta x_n$$

will give us an estimate for the total length of the curve.

The length of
a straight line



The length of
a curved line

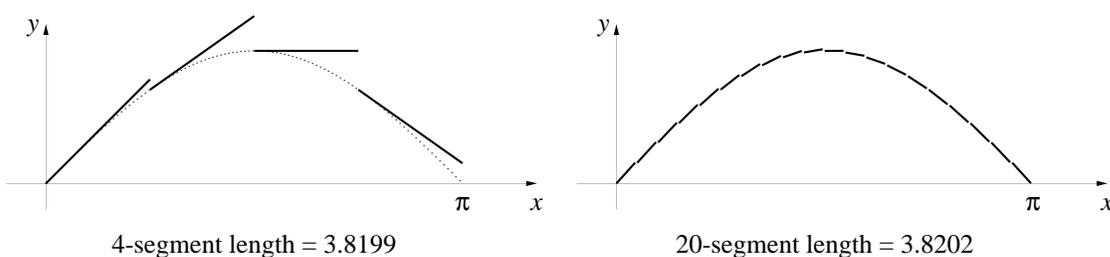


The length of a curve is estimated by a Riemann sum

Once again, we find an expression that has the form of a Riemann sum. There is, however, a new ingredient worth noting. The estimate is a Riemann sum not for the *original* function $g(x)$ but for *another* function

$$f(x) = \sqrt{1 + (g'(x))^2}$$

that we constructed using g . The important thing is that the length is estimated by a Riemann sum for *some* function.



The figure above shows two estimates for the length of the graph of $y = \sin x$ between 0 and π . In each case, we used subintervals of equal length and we sampled the slope at the left end of each subinterval. As you can see, the four segments approximate the graph of $y = \sin x$ only very roughly. When we increase the number of segments to 20, on the right, the approximation to the shape of the graph becomes quite good. Notice that the graph itself is not shown on the right; only the 20 segments.

To calculate the two lengths, we constructed Riemann sums for the function $f(x) = \sqrt{1 + \cos^2 x}$. We used the fact that the derivative of $g(x) = \sin x$ is $g'(x) = \cos x$, and we did the calculations using the program RIEMANN. By using the program with still smaller subintervals you can show that

$$\text{the exact length} = 3.820197789 \dots$$

Thus, the 20-segment estimate is already accurate to four decimal places.

We have already constructed estimates for the length of a curve, in chapter 2 (pages 89–91). Those estimates were sums, too, but they were not *Riemann* sums. The terms had the form $\sqrt{\Delta x^2 + \Delta y^2}$; they were not *products* of the form $\sqrt{1 + m^2} \cdot \Delta x$. The sums in chapter 2 may seem more straightforward. However, we are developing Riemann sums as a powerful general tool for dealing with many different questions. By expressing lengths as Riemann sums we gain access to that power.

Definition

The Riemann sums that appear in the calculation of power, distance, and length are instances of a general mathematical object that can be constructed for *any* function whatsoever. We pause now to describe that construction apart from any particular context. In what follows it will be convenient for us to write an interval of the form $a \leq x \leq b$ more compactly as $[a, b]$.

Notation: $[a, b]$

Definition. Suppose the function $f(x)$ is defined for x in the interval $[a, b]$. Then a **Riemann sum** for $f(x)$ on $[a, b]$ is an expression of the form

$$f(x_1) \cdot \Delta x_1 + f(x_2) \cdot \Delta x_2 + \cdots + f(x_n) \cdot \Delta x_n.$$

The interval $[a, b]$ has been divided into n subintervals whose lengths are $\Delta x_1, \dots, \Delta x_n$, respectively, and for each k from 1 to n , x_k is some point in the k -th subinterval.

Notice that once the function and the interval have been specified, a Riemann sum is determined by the following data:

Data for a Riemann sum

- A **decomposition** of the original interval into subintervals (which determines the lengths of the subintervals).
- A **sampling point** chosen from each subinterval (which determines a value of the function on each subinterval).

A Riemann sum for $f(x)$ is a sum of products of values of Δx and values of $y = f(x)$. If x and y have units, then so does the Riemann sum; its units are the units for x times the units for y . When a Riemann sum arises in a particular context, the notation may look different from what appears in the definition just given: the variable might not be x , and the function might not be $f(x)$. For example, the energy approximation we considered at the beginning of the section is a Riemann sum for the power demand function $p(t)$ on $[0, 24]$. The length approximation for the graph $y = \sin x$ is a Riemann sum for the function $\sqrt{1 + \cos^2 x}$ on $[0, \pi]$.

Units

It is important to note that, from a mathematical point of view, a Riemann sum is just a number. It's the *context* that provides the meaning: Riemann sums for a power demand that varies over time approximate total

A Riemann sum is just a number

energy consumption; Riemann sums for a velocity that varies over time approximate total distance; and Riemann sums for a height that varies over distance approximate total area.

To illustrate the generality of a Riemann sum, and to stress that it is just a number arrived at through arbitrary choices, let's work through an example without a context. Consider the function

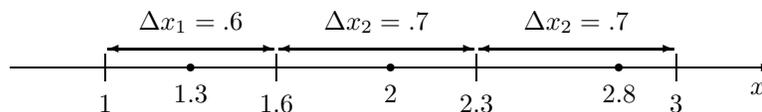
$$f(x) = \sqrt{1+x^3} \quad \text{on} \quad [1, 3].$$

The data

We will break up the full interval $[1, 3]$ into three subintervals $[1, 1.6]$, $[1.6, 2.3]$ and $[2.3, 3]$. Thus

$$\Delta x_1 = .6 \quad \Delta x_2 = \Delta x_3 = .7.$$

Next we'll pick a point in each subinterval, say $x_1 = 1.3$, $x_2 = 2$ and $x_3 = 2.8$. Here is the data laid out on the x -axis.



With this data we get the following Riemann sum for $\sqrt{1+x^3}$ on $[1, 3]$:

$$\begin{aligned} & f(x_1) \cdot \Delta x_1 + f(x_2) \cdot \Delta x_2 + f(x_3) \cdot \Delta x_3 \\ &= \sqrt{1+1.3^3} \times .6 + \sqrt{1+2^3} \times .7 + \sqrt{1+2.8^3} \times .7 \\ &= 6.5263866 \end{aligned}$$

In this case, the choice of the subintervals, as well as the choice of the point x_k in each subinterval, was haphazard. Different data would produce a different value for the Riemann sum.

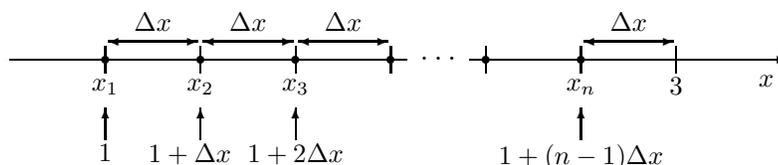
Keep in mind that an individual Riemann sum is not especially significant. Ultimately, we are interested in seeing what happens when we recalculate Riemann sums with smaller and smaller subintervals. For that reason, it is helpful to do the calculations systematically.

Systematic data

Calculating a Riemann sum algorithmically. As we have seen with our contextual problems, the data for a Riemann sum is not usually chosen in a haphazard fashion. In fact, when dealing with functions given by formulas, such as the function $f(x) = \sqrt{1-x^2}$ whose graph is a semicircle, it pays to be systematic. We use subintervals of equal size and pick the “same” point

from each subinterval (e.g., always pick the midpoint or always pick the left endpoint). The benefit of systematic choices is that we can write down the computations involved in a Riemann sum in a simple algorithmic form that can be carried out on a computer.

Let's illustrate how this strategy applies to the function $\sqrt{1+x^3}$ on $[1, 3]$. Since the whole interval is $3 - 1 = 2$ units long, if we construct n subintervals of equal length Δx , then $\Delta x = 2/n$. For every $k = 1, \dots, n$, we choose the sampling point x_k to be the left endpoint of the k -th subinterval. Here is a picture of the data:



In this systematic approach, the space between one sampling point and the next is Δx , the same as the width of a subinterval. This puts the k -th sampling point at $x = 1 + (k - 1)\Delta x$.

In the following table, we add up the terms in a Riemann sum S for $f(x) = \sqrt{1+x^3}$ on the interval $[1, 3]$. We used $n = 4$ subintervals and always sampled f at the left endpoint. Each row shows the following:

1. the current sampling point;
2. the value of f at that point;
3. the current term $\Delta S = f \cdot \Delta x$ in the sum;
4. the accumulated value of S .

left endpoint	current $\sqrt{1+x^3}$	current ΔS	accumulated S
1	1.4142	.7071	.7071
1.5	2.0917	1.0458	1.7529
2	3	1.5	3.2529
2.5	4.0774	2.0387	5.2916

The Riemann sum S appears as the final value 5.2916 in the fourth column.

The program RIEMANN, below, will generate the last two columns in the above table. The statement `x = a` on the sixth line determines the position

of the first sampling point. Within the FOR–NEXT loop, the statement `x = x + deltax` moves the sampling point to its next position.

Program: RIEMANN
Left endpoint Riemann sums

```

DEF fnf (x) = SQR(1 + x ^ 3)
a = 1
b = 3
numberofsteps = 4
deltax = (b - a) / numberofsteps
x = a
accumulation = 0
FOR k = 1 TO numberofsteps
    deltaS = fnf(x) * deltax
    accumulation = accumulation + deltaS
    x = x + deltax
    PRINT deltaS, accumulation
NEXT k

```

By modifying RIEMANN, you can calculate Riemann sums for other sampling points and for other functions. For example, to sample at midpoints, you must start at the midpoint of the first subinterval. Since the subinterval is Δx units wide, its midpoint is $\Delta x/2$ units from the left endpoint, which is $x = a$. Thus, to have the program generate midpoint sums, just change the statement that “initializes” x (line 6) to `x = a + deltax / 2`.

Summation Notation

Because Riemann sums arise frequently and because they are unwieldy to write out in full, we now introduce a method—called **summation notation**—that allows us to write them more compactly. To see how it works, look first at the sum

$$1^2 + 2^2 + 3^2 + \cdots + 50^2.$$

Using summation notation, we can express this as

$$\sum_{k=1}^{50} k^2.$$

For a somewhat more abstract example, consider the sum

$$a_1 + a_2 + a_3 + \cdots + a_n,$$

which we can express as

$$\sum_{k=1}^n a_k.$$

We use the capital letter *sigma* \sum from the Greek alphabet to denote a sum. For this reason, summation notation is sometimes referred to as **sigma notation**. You should regard \sum as an instruction telling you to **sum** the numbers of the indicated form as the index k runs through the integers, starting at the integer displayed below the \sum and ending at the integer displayed above it. Notice that changing the index k to some other letter has no effect on the sum. For example,

Sigma notation

$$\sum_{k=1}^{20} k^3 = \sum_{j=1}^{20} j^3,$$

since both expressions give the sum of the cubes of the first twenty positive integers. Other aspects of summation notation will be covered in the exercises.

Summation notation allows us to write the Riemann sum

$$f(x_1) \cdot \Delta x_1 + \cdots + f(x_n) \cdot \Delta x_n$$

more efficiently as

$$\sum_{k=1}^n f(x_k) \cdot \Delta x_k.$$

Be sure not to get tied into one particular way of using these symbols. For example, you should instantly recognize

$$\sum_{i=1}^m \Delta t_i g(t_i)$$

as a Riemann sum. In what follows we will commonly use summation notation when working with Riemann sums. The important thing to remember is that summation notation is only a “shorthand” to express a Riemann sum in a more compact form.

Exercises

Making approximations

1. Estimate the average velocity of the ship whose motion is described on page 352. The voyage lasts 15 hours.
2. The aim of this question is to determine how much electrical energy was consumed in a house over a 24-hour period, when the power demand p was measured at different times to have these values:

time (24-hour clock)	power (watts)
1:30	275
5:00	240
8:00	730
9:30	300
11:00	150
15:00	225
18:30	1880
20:00	950
22:30	700
23:00	350

Notice that the time interval is from $t = 0$ hours to $t = 24$ hours, but the power demand was not sampled at either of those times.

- a) Set up an estimate for the energy consumption in the form of a Riemann sum $p(t_1)\Delta t_1 + \cdots + p(t_n)\Delta t_n$ for the power function $p(t)$. To do this, you must identify explicitly the value of n , the sampling times t_k , and the time intervals Δt_k that you used in constructing your estimate. [Note: the sampling times come from the table, but there is wide latitude in how you choose the subintervals Δt_k .]
- b) What is the estimated energy consumption, using your choice of data? There is no single “correct” answer to this question. Your estimate depends on the choices you made in setting up the Riemann sum.
- c) Plot the data given in the table in part (a) on a (t, p) -coordinate plane. Then draw on the same coordinate plane the step function that represents your estimate of the power function $p(t)$. The width of the k -th step should be the time interval Δt_k that you specified in part (a); is it?

d) Estimate the *average* power demand in the house during the 24-hour period.

Waste production. A colony of living yeast cells in a vat of fermenting grape juice produces waste products—mainly alcohol and carbon dioxide—as it consumes the sugar in the grape juice. It is reasonable to expect that another yeast colony, twice as large as this one, would produce twice as much waste over the same time period. Moreover, since waste accumulates over time, if we double the time period we would expect our colony to produce twice as much waste.

These observations suggest that waste production is proportional to both the size of the colony and the amount of time that passes. If P is the size of the colony, in grams, and Δt is a short time interval, then we can express waste production W as a function of P and Δt :

$$W = k \cdot P \cdot \Delta t \text{ grams.}$$

If Δt is measured in hours, then the multiplier k has to be measured in units of grams of waste per hour per gram of yeast.

The preceding formula is useful only over a time interval Δt in which the population size P does not vary significantly. If the time interval is large, and the population size can be expressed as a function $P(t)$ of the time t , then we can estimate waste production by breaking up the whole time interval into a succession of smaller intervals $\Delta t_1, \Delta t_2, \dots, \Delta t_n$ and forming a Riemann sum

$$k P(t_1) \Delta t_1 + \dots + k P(t_n) \Delta t_n \approx W \text{ grams.}$$

The time t_k must lie within the time interval Δt_k , and $P(t_k)$ must be a good approximation to the population size $P(t)$ throughout that time interval.

3. Suppose the colony starts with 300 grams of yeast (i.e., at time $t = 0$ hours) and it grows exponentially according to the formula

$$P(t) = 300 e^{0.2t}.$$

If the waste production constant k is 0.1 grams per hour per gram of yeast, estimate how much waste is produced in the first four hours. Use a Riemann sum with four hour-long time intervals and measure the population size of the yeast in the middle of each interval—that is, “on the half-hour.”

Using RIEMANN

4. a) Calculate left endpoint Riemann sums for the function $\sqrt{1+x^3}$ on the interval $[1, 3]$ using 40, 400, 4000, and 40000 equally-spaced subintervals. How many decimal points in this sequence have stabilized?
- b) The left endpoint Riemann sums for $\sqrt{1+x^3}$ on the interval $[1, 3]$ seem to be approaching a limit as the number of subintervals increases without bound. Give the numerical value of that limit, accurate to four decimal places.
- c) Calculate left endpoint Riemann sums for the function $\sqrt{1+x^3}$ on the interval $[3, 7]$. Construct a sequence of Riemann sums using more and more subintervals, until you can determine the limiting value of these sums, accurate to three decimal places. What is that limit?
- d) Calculate left endpoint Riemann sums for the function $\sqrt{1+x^3}$ on the interval $[1, 7]$ in order to determine the limiting value of the sums to three decimal place accuracy. What is that value? How are the limiting values in parts (b), (c), and (d) related? How are the corresponding *intervals* related?
5. Modify RIEMANN so it will calculate a Riemann sum by sampling the given function at the *midpoint* of each subinterval, instead of the left endpoint. Describe exactly how you changed the program to do this.
6. a) Calculate *midpoint* Riemann sums for the function $\sqrt{1+x^3}$ on the interval $[1, 3]$ using 40, 400, 4000, and 40000 equally-spaced subintervals. How many decimal points in this sequence have stabilized?
- b) Roughly how many subintervals are needed to make the midpoint Riemann sums for $\sqrt{1+x^3}$ on the interval $[1, 3]$ stabilize out to the first four digits? What is the stable value? Compare this to the limiting value you found earlier for left endpoint Riemann sums. Is one value larger than the other; could they be the same?
- c) Comment on the relative “efficiency” of midpoint Riemann sums versus left endpoint Riemann sums (at least for the function $\sqrt{1+x^3}$ on the interval $[1, 3]$). To get the same level of accuracy, an *efficient* calculation will take fewer steps than an *inefficient* one.
7. a) Modify RIEMANN to calculate *right endpoint* Riemann sums, and use it to calculate right endpoint Riemann sums for the function $\sqrt{1+x^3}$ on

the interval $[1, 3]$ using 40, 400, 4000, and 40000 equally-spaced subintervals. How many digits in this sequence have stabilized?

b) Comment on the efficiency of right endpoint Riemann sums as compared to left endpoint and to midpoint Riemann sums—at least as far as the function $\sqrt{1+x^3}$ is concerned.

8. Calculate left endpoint Riemann sums for the function

$$f(x) = \sqrt{1-x^2} \quad \text{on the interval } [-1, 1].$$

Use 20 and 50 equally-spaced subintervals. Compare your values with the estimates for the area of a semicircle given on page 356.

9. a) Calculate left endpoint Riemann sums for the function

$$f(x) = \sqrt{1+\cos^2 x} \quad \text{on the interval } [0, \pi].$$

Use 4 and 20 equally-spaced subintervals. Compare your values with the estimates for the length of the graph of $y = \sin x$ between 0 and π , given on page 358.

b) What is the limiting value of the Riemann sums, as the number of subintervals becomes infinite? Find the limit to 11 decimal places accuracy.

10. Calculate left endpoint Riemann sums for the function

$$f(x) = \cos(x^2) \quad \text{on the interval } [0, 4],$$

using 100, 1000, and 10000 equally-spaced subintervals.

[Answer: With 10000 equally-spaced intervals, the left endpoint Riemann sum has the value .59485189.]

11. Calculate left endpoint Riemann sums for the function

$$f(x) = \frac{\cos x}{1+x^2} \quad \text{on the interval } [2, 3],$$

using 10, 100, and 1000 equally-spaced subintervals. The Riemann sums are all negative; why? (A suggestion: sketch the graph of f . What does that tell you about the signs of the terms in a Riemann sum for f ?)

12. a) Calculate midpoint Riemann sums for the function

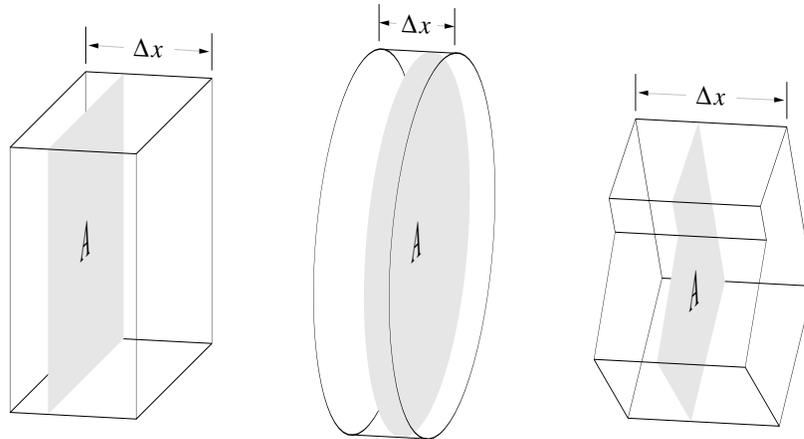
$$H(z) = z^3 \quad \text{on the interval } [-2, 2],$$

using 10, 100, and 1000 equally-spaced subintervals. The Riemann sums are all zero; why? (On some computers and calculators, you may find that there will be a nonzero digit in the fourteenth or fifteenth decimal place – this is due to “round-off error”.)

- b) Repeat part (a) using *left endpoint* Riemann sums. Are the results still zero? Can you explain the difference, if any, between these two results?

Volume as a Riemann sum

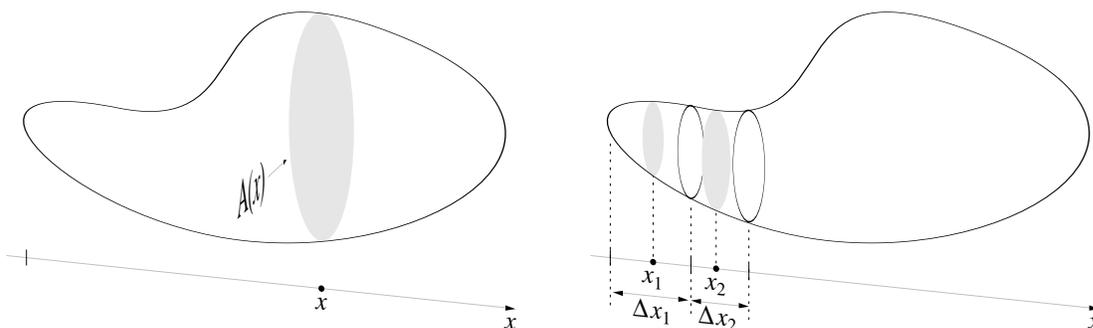
If you slice a rectangular parallelepiped (e.g., a brick or a shoebox) parallel to a face, the area A of a **cross-section** does not vary. The same is true for a cylinder (e.g., a can of spinach or a coin). For *any* solid that has a constant cross-section (e.g., the object on the right, below), its volume is just the product of its cross-sectional area with its thickness.



$$\text{volume} = \text{area of cross-section} \times \text{thickness} = A \cdot \Delta x$$

Most solids don't have such a regular shape. They are more like the one shown below. If you take cross-sectional slices perpendicular to some fixed line (which will become our x -axis), the slices will not generally have a regular shape. They may be roughly oval, as shown below, but they will generally vary in area. Suppose the area of the cross-section x inches along the axis is $A(x)$ square inches. Because $A(x)$ varies with x , you cannot calculate

the volume of this solid using the simple formula above. However, you can *estimate* the volume as a Riemann sum for A .



The procedure should now be familiar to you. Subdivide the x -axis into segments of length $\Delta x_1, \Delta x_2, \dots, \Delta x_n$ inches, respectively. The solid piece that lies over the first segment has a thickness of Δx_1 inches. If you slice this piece at a point x_1 inches along the x -axis, the area of the slice is $A(x_1)$ square inches, and the volume of the piece is approximately $A(x_1) \cdot \Delta x_1$ cubic inches. The second piece is Δx_2 inches thick. If you slice it x_2 inches along the x -axis, the slice has an area of $A(x_2)$ square inches, so the second piece has an approximate volume of $A(x_2) \cdot \Delta x_2$ cubic inches. If you continue in this way and add up the n volumes, you get an estimate for the total volume that has the form of a Riemann sum for the area function $A(x)$:

$$\text{volume} \approx A(x_1) \Delta x_1 + A(x_2) \Delta x_2 + \cdots + A(x_n) \Delta x_n \text{ cubic inches.}$$

One place where this approach can be used is in medical diagnosis. The X-ray technique known as a CAT scan provides a sequence of precisely-spaced cross-sectional views of a patient. From these views much information about the state of the patient's internal organs can be gained without invasive surgery. In particular, the volume of a specific piece of tissue can be estimated, as a Riemann sum, from the areas of individual slices and the spacing between them. The next exercise gives an example.

13. A CAT scan of a human liver shows us X-ray “slices” spaced 2 centimeters apart. If the areas of the slices are 72, 145, 139, 127, 111, 89, 63, and 22 square centimeters, estimate the volume of the liver.

14. The volume of a sphere whose radius is r is exactly $V = 4\pi r^3/3$.

- a) Using the formula, determine the volume of the sphere whose radius is 3. Give the numerical value to four decimal places accuracy.

One way to get a sphere of radius 3 is to rotate the graph of the semi-circle

$$r(x) = \sqrt{9 - x^2} \quad -3 \leq x \leq 3$$

around the x -axis. Every cross-section perpendicular to the x -axis is a circle. At the point x , the radius of the circle is $r(x)$, and its area is

$$A = \pi r^2 = \pi(r(x))^2 = A(x).$$

You can thus get estimates for the volume of the sphere by constructing Riemann sums for $A(x)$ on the interval $[-3, 3]$.

- b) Calculate a sequence of estimates for the volume of the sphere that use more and more slices, until the value of the estimate stabilizes out to four decimal places. Does this value agree with the value given by the formula in part (a)?

15. a) Rotate the graph of $r(x) = .5x$, with $0 \leq x \leq 6$ around the x -axis. What shape do you get? Describe it precisely, and find its volume using an appropriate geometric formula.

- b) Calculate a sequence of estimates for the volume of the same object by constructing Riemann sums for the area function $A(x) = \pi(r(x))^2$. Continue until your estimates stabilize out to four decimal places. What value do you get?

Summation notation

16. Determine the numerical value of each of the following:

$$\text{a) } \sum_{k=1}^{10} k \qquad \text{b) } \sum_{k=1}^5 k^2 \qquad \text{c) } \sum_{j=0}^4 2j + 1$$

[Answer: $\sum_{k=1}^5 k^2 = 55.$]

17. Write “the sum of the first five positive even integers” in summation notation.

18. Determine the numerical value of

$$\text{a) } \sum_{n=1}^5 \left(\frac{1}{n} - \frac{1}{n+1} \right) \qquad \text{b) } \sum_{n=1}^{500} \left(\frac{1}{n} - \frac{1}{n+1} \right)$$

19. Express the following sums using summation notation.

- a) $1^2 + 2^2 + 3^2 + \cdots + n^2$.
 b) $2^1 + 2^2 + 2^3 + \cdots + 2^m$.
 c) $f(s_1)\Delta s + f(s_2)\Delta s + \cdots + f(s_{12})\Delta s$.
 d) $y_1^2 \Delta y_1 + y_2^2 \Delta y_2 + \cdots + y_n^2 \Delta y_n$.

20. Express each of the following as a sum written out term-by-term. (There is no need to calculate the numerical value, even when that can be done.)

$$\text{a) } \sum_{l=3}^{n-1} a_l \qquad \text{b) } \sum_{j=0}^4 \frac{j+1}{j^2+1} \qquad \text{c) } \sum_{k=1}^5 H(x_k) \Delta x_k$$

21. Acquire experimental evidence for the claim

$$\left(\sum_{k=1}^n k \right)^2 = \sum_{k=1}^n k^3$$

by determining the numerical values of both sides of the equation for $n = 2, 3, 4, 5,$ and 6 .

22. Let $g(u) = 25 - u^2$ and suppose the interval $[0, 2]$ has been divided into 4 equal subintervals Δu and u_j is the left endpoint of the j -th interval. Determine the numerical value of the Riemann sum

$$\sum_{j=1}^4 g(u_j) \Delta u.$$

Length and area

23. Using Riemann sums with equal subintervals, estimate the length of the parabola $y = x^2$ over the interval $0 \leq x \leq 1$. Obtain a sequence of estimates that stabilize to four decimal places. How many subintervals did you need? (Compare your result here with the earlier result on page 96.)

24. Using Riemann sums, obtain a sequence of estimates for the area under each of the following curves. Continue until the first four decimal places stabilize in your estimates.

a) $y = x^2$ over $[0, 1]$ b) $y = x^2$ over $[0, 3]$ c) $y = x \sin x$ over $[0, \pi]$

25. What is the area under the curve $y = \exp(-x^2)$ over the interval $[0, 1]$? Give an estimate that is accurate to four decimal places. Sketch the curve and shade the area.

26. a) Estimate, to four decimal place accuracy, the length of the graph of the natural logarithm function $y = \ln x$ over the interval $[1, e]$.

b) Estimate, to four decimal place accuracy, the length of the graph of the exponential function $y = \exp(x)$ over the interval $[0, 1]$.

27. a) What is the length of the hyperbola $y = 1/x$ over the interval $[1, 4]$? Obtain an estimate that is accurate to four decimal places.

b) What is the area under the hyperbola over the same interval? Obtain an estimate that is accurate to four decimal places.

28. The graph of $y = \sqrt{4 - x^2}$ is a semicircle whose radius is 2. The circumference of the whole circle is 4π , so the length of the part of the circle in the first quadrant is exactly π .

a) Using left endpoint Riemann sums, estimate the length of the graph $y = \sqrt{4 - x^2}$ over the interval $[0, 2]$ in the first quadrant. How many subintervals did you need in order to get an estimate that has the value 3.14159...?

b) There is a technical problem that makes it impossible to use *right* endpoint Riemann sums. What is the problem?

6.3 The Integral

Refining Riemann Sums

In the last section, we estimated the electrical energy a town consumed by constructing a Riemann sum for the power demand function $p(t)$. Because we sampled the power function only five times in a 24-hour period, our estimate was fairly rough. We would get a better estimate by sampling more frequently—that is, by constructing a Riemann sum with more terms and shorter subintervals. The process of refining Riemann sums in this way leads to the mathematical object called the **integral**.

To see what an integral is—and how it emerges from this process of refining Riemann sums—let’s return to the function

$$\sqrt{1+x^3} \quad \text{on} \quad [1, 3]$$

we analyzed at the end of the last section. What happens when we refine Riemann sums for this function by using smaller subintervals? If we systematically choose n equal subintervals and evaluate $\sqrt{1+x^3}$ at the left endpoint of each subinterval, then we can use the program RIEMANN (page 362) to produce the values in the following table. For future reference we record the size of the subinterval $\Delta x = 2/n$ as well.

Refining Riemann sums
with equal subintervals

Left endpoint Riemann sums for $\sqrt{1+x^3}$ on $[1, 3]$

n	Δx	Riemann sum
100	.02	6.191 236 2
1 000	.002	6.226 082 6
10 000	.0002	6.229 571 7
100 000	.00002	6.229 920 6

The first four digits have stabilized, suggesting that these Riemann sums, at least, approach the limit 6.229. . . .

It’s too soon to say that *all* the Riemann sums for $\sqrt{1+x^3}$ on the interval $[1, 3]$ approach this limit, though. There is such an enormous diversity of choices at our disposal when we construct a Riemann sum. We haven’t seen what happens, for instance, if we choose midpoints instead of left endpoints, or if we choose subintervals that are not all of the same size. Let’s explore the first possibility.

Midpoints versus
left endpoints

To modify RIEMANN to choose midpoints, we need only change the line

$$x = a$$

that determines the position of the first sampling point to

$$x = a + \text{deltax} / 2.$$

With this modification, RIEMANN produces the following data.

Midpoint Riemann sums for $\sqrt{1+x^3}$ on $[1, 3]$

n	Δx	Riemann sum
10	.2	6.227 476 5
100	.02	6.229 934 5
1 000	.002	6.229 959 1
10 000	.0002	6.229 959 4

This time, the first *seven* digits have stabilized, even though we used only 10,000 subintervals—ten times fewer than we needed to get four digits to stabilize using left endpoints! This is further evidence that the Riemann sums converge to a limit, and we can even specify the limit more precisely as 6.229959. . . .

These tables also suggest that midpoints are more “efficient” than left endpoints in revealing the limiting value of successive Riemann sums. This is indeed true. In Chapter 11, we will look into this further.

We still have another possibility to consider: what happens if we choose subintervals of different sizes—as we did in calculating the energy consumption of a town? By allowing variable subintervals, we make the problem messier to deal with, but it does not become conceptually more difficult. In fact, we can still get all the information we really need in order to understand what happens when we refine Riemann sums.

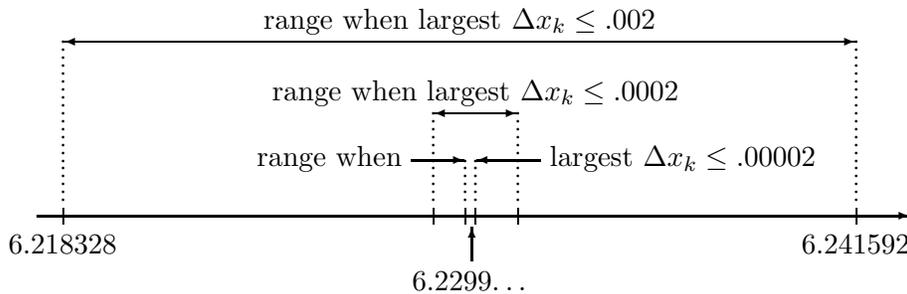
To see what form this information will take, look back at the two tables for midpoints and left endpoints, and compare the values of the Riemann sums that they report for the same size subinterval. When the subinterval was fairly large, the values differed by a relatively large amount. For example, when $\Delta x = .02$ we got two sums (namely 6.2299345 and 6.1912362) that differ by more than .038. As the subinterval got smaller, the difference between the Riemann sums got smaller too. (When $\Delta x = .0002$, the sums differ by

less than .0004.) This is the general pattern. That is to say, for subintervals with a given maximum size, the various Riemann sums that can be produced will still differ from one another, but those sums will all lie within a certain range that gets smaller as the size of the largest subinterval gets smaller.

The connection between the range of Riemann sums and the size of the largest subinterval is subtle and technically complex; this course will not explore it in detail. However, we can at least see what happens concretely to Riemann sums for the function $\sqrt{1+x^3}$ over the interval $[1, 3]$. The following table shows the smallest and largest possible Riemann sum that can be produced when no subinterval is larger than the maximum size Δx_k given in the first column.

Refining Riemann sums with *unequal* subintervals

maximum size of Δx_k	Riemann sums range from	to	difference between extremes
.02	6.113 690	6.346 328	.232 638
.002	6.218 328	6.241 592	.023 264
.000 2	6.228 796	6.231 122	.002 326
.000 02	6.229 843	6.230 076	.000 233
.000 002	6.229 948	6.229 971	.000 023



The range of Riemann sums for $\sqrt{1+x^3}$ on $[1, 3]$

This table provides the most compelling evidence that there is a single number 6.2299... that *all* Riemann sums will be arbitrarily close to, if they are constructed with sufficiently small subintervals Δx_k . This number is called the **integral** of $\sqrt{1+x^3}$ on the interval $[1, 3]$, and we will express this by writing

The integral

$$\int_1^3 \sqrt{1+x^3} dx = 6.2299\dots$$

Each Riemann sum approximates this integral, and in general the approximations get better as the size of the largest subinterval is made smaller.

Moreover, as the subintervals get smaller, the location of the sampling points matters less and less.

The unusual symbol \int that appears here reflects the historical origins of the integral. We'll have more to say about it after we consider the definition.

Definition

The purpose of the following definition is to give a name to the number to which the Riemann sums for a function converge, *when those sums do indeed converge*.

Definition. Suppose all the Riemann sums for a function $f(x)$ on an interval $[a, b]$ get arbitrarily close to a single number when the lengths $\Delta x_1, \dots, \Delta x_n$ are made small enough. Then this number is called the **integral** of $f(x)$ on $[a, b]$ and it is denoted

$$\int_a^b f(x) dx.$$

The function f is called the **integrand**. The definition begins with a *Suppose ...* because there are functions whose Riemann sums don't converge. We'll look at an example on page 378. However, that example is quite special. All the functions that typically arise in context, and nearly all the functions we study in calculus, *do* have integrals. In particular, every continuous function has an integral, and so do many non-continuous functions—such as the step functions with which we began this chapter. (Continuous functions are discussed on pages 304–307).

Notice that the definition doesn't speak about the choice of sampling points. The condition that the Riemann sums be close to a single number involves only the subintervals $\Delta x_1, \Delta x_2, \dots, \Delta x_n$. This is important; it says *once the subintervals are small enough, it doesn't matter which sampling points x_k we choose—all of the Riemann sums will be close to the value of the integral*. (Of course, some will still be closer to the value of the integral than others.)

The integral allows us to resolve the dilemma we stated at the beginning of the chapter. Here is the dilemma: how can we describe the product of two quantities when one of them varies? Consider, for example, how we expressed the energy consumption of a town over a 24-hour period. The basic relation

A typical function has an integral

How an integral expresses a product

$$\text{energy} = \text{power} \times \text{elapsed time}$$

cannot be used directly, because power demand varies. Indirectly, though, we can use the relation to build a Riemann sum for power demand p over time. This gives us an *approximation*:

$$\text{energy} \approx \sum_{k=1}^n p(t_k) \Delta t_k \quad \text{megawatt-hours.}$$

As these sums are refined, two things happen. First, they converge to the true level of energy consumption. Second, they converge to the integral—by the definition of the integral. Thus, energy consumption is described *exactly* by the integral

$$\text{energy} = \int_0^{24} p(t) dt \quad \text{megawatt-hours}$$

of the power demand p . In other words, *energy is the integral of power over time*.

On page 352 we asked how far a ship would travel in 15 hours if we knew its velocity was $v(t)$ miles per hour at time t . We saw the distance could be estimated by a Riemann sum for the v . Therefore, reasoning just as we did for energy, we conclude that the *exact* distance is given by the integral

$$\text{distance} = \int_0^{15} v(t) dt \quad \text{miles.}$$

The energy integral has the same units as the Riemann sums that approximate it. Its units are the product of the megawatts used to measure p and the hours used to measure dt (or t). The units for the distance integral are the product of the miles per hour used to measure velocity and the hours used to measure time. In general, the units for the integral

The units for
an integral

$$\int_a^b f(x) dx$$

are the product of the units for f and the units for x .

Because the integral is approximated by its Riemann sums, we can use summation notation (introduced in the previous section) to write

Why \int is used
to denote an integral

$$\int_1^3 \sqrt{1+x^3} dx \approx \sum_{k=1}^n \sqrt{1+x_k^3} \Delta x_k.$$

This expression helps reveal where the rather unusual-looking notation for the integral comes from. In seventeenth century Europe (when calculus was being created), the letter ‘s’ was written two ways: as ‘s’ and as ‘ \int ’. The \int that appears in the integral and the \sum that appears in the Riemann sum both serve as abbreviations for the word *sum*. While we think of the Riemann sum as a sum of products of the form $\sqrt{1+x_k^3} \cdot \Delta x_k$, in which the various Δx_k are small quantities, some of the early users of calculus thought of the integral as a sum of products of the form $\sqrt{1+x^3} \cdot dx$, in which dx is an “infinitesimally” small quantity.

Now we do not use infinitesimals or regard the integral as a sum directly. On the contrary, for us the integral is a *limit* of Riemann sums as the subinterval lengths Δx_k all shrink to 0. In fact, we can express the integral directly as a limit:

$$\int_a^b f(x) dx = \lim_{\Delta x_k \rightarrow 0} \sum_{k=1}^n f(x_k) \Delta x_k.$$

Integration
“puts together”
products of the form
 $f(x) \Delta x$

The process of calculating an integral is called **integration**. Integration means “putting together.” To see why this name is appropriate, notice that we determine energy consumption over a long time interval by putting together a lot of energy computations $p \cdot \Delta t$ over a succession of short periods.

A function that does not have an integral

Riemann sums converge to a single number for many functions—but not all. For example, Riemann sums for

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

do not converge. Let’s see why.

A rational number is the quotient p/q of one integer p by another q . An irrational number is one that is *not* such a quotient; for example, $\sqrt{2}$ is irrational. The values of J are continually jumping between 0 and 1.

Why Riemann sums for
 $J(x)$ do not converge

Suppose we construct a Riemann sum for J on the interval $[0, 1]$ using the subintervals $\Delta x_1, \Delta x_2, \dots, \Delta x_n$. Every subinterval contains both rational and irrational numbers. Thus, we could choose all the sampling points to be rational numbers r_1, r_2, \dots, r_n . In that case,

$$J(r_1) = J(r_2) = \dots = J(r_n) = 0,$$

so the Riemann sum has the value

$$\sum_{k=1}^n J(r_k) \Delta x_k = \sum_{k=1}^n 0 \cdot \Delta x_k = 0.$$

But we could also choose all the sampling points to be irrational numbers s_1, s_2, \dots, s_n . In that case,

$$J(s_1) = J(s_2) = \dots = J(s_n) = 1,$$

and the Riemann sum would have the value

$$\sum_{k=1}^n J(s_k) \Delta x_k = \sum_{k=1}^n 1 \cdot \Delta x_k = \sum_{k=1}^n \Delta x_k = 1.$$

(For any subdivision Δx_k , $\Delta x_1 + \Delta x_2 + \dots + \Delta x_n = 1$ because the subintervals together form the interval $[0, 1]$.) If some sampling points are rational and others irrational, the value of the Riemann sum will lie somewhere between 0 and 1. Thus, the Riemann sums range from 0 to 1, *no matter how small the subintervals Δx_k are chosen*. They cannot converge to any single number.

The function J shows us that not every function has an integral. The definition of the integral (page 376) takes this into account. It doesn't guarantee that an arbitrary function will have an integral. It simply says that *if* the Riemann sums converge to a single number, then we can give that number a name—the integral.

Can you imagine what the graph of $y = J(x)$ would look like? It would consist of two horizontal lines with gaps; in the upper line, the gaps would be at all the rational points, in the lower at the irrational points. This is impossible to draw! Roughly speaking, any graph you can draw you can integrate.

Visualizing the Integral

The eye plays an important role in our thinking. We visualize concepts whenever we can. Given a function $y = f(x)$, we visualize it as a graph in the x, y -coordinate plane. We visualize the derivative of f as the slope of its graph at any point. We can visualize the integral of f , too. We will view it as the area under that graph. Let's see why we can.

We have already made a connection between areas and Riemann sums, in the last section. Our starting point was the basic formula

area = height \times width.

Since the height of the graph $y = f(x)$ at any point x is just $f(x)$, we were tempted to say that

$$\text{area} = f(x) \cdot (b - a).$$

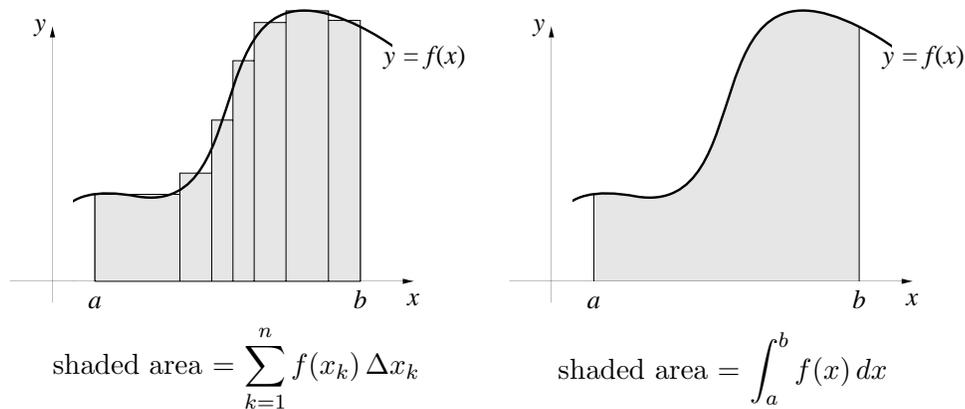
Of course we couldn't do this, because the height $f(x)$ is variable. The remedy was to slice up the interval $[a, b]$ into small pieces Δx_k , and assemble a collection of products $f(x_k) \Delta x_k$:

$$\sum_{k=1}^n f(x_k) \Delta x_k = f(x_1) \Delta x_1 + f(x_2) \Delta x_2 + \cdots + f(x_n) \Delta x_n.$$

Riemann sums
converge to the
area ...

... and to the integral

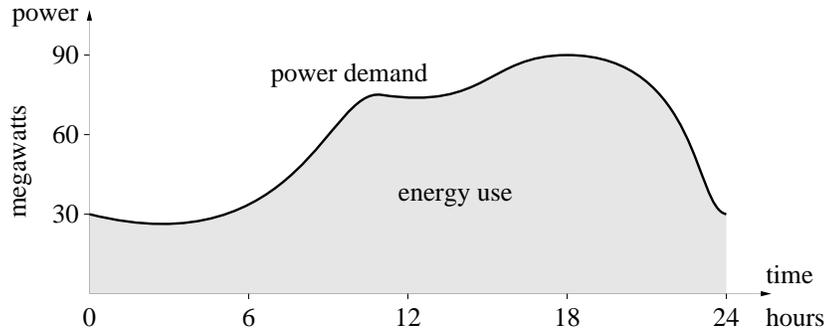
This is a Riemann sum. It represents the total area of a row of side-by-side rectangles whose tops approximate the graph of f . As the Riemann sums are refined, the tops of the rectangles approach the shape of the graph, and their areas approach the area under the graph. But the process of refining Riemann sums leads to the integral, so the integral must be the area under the graph.



Every integral we have encountered can be visualized as the area under a graph. For instance, since

$$\text{energy use} = \int_0^{24} p(t) dt,$$

we can now say that the energy used by a town is just the area under its power demand graph.



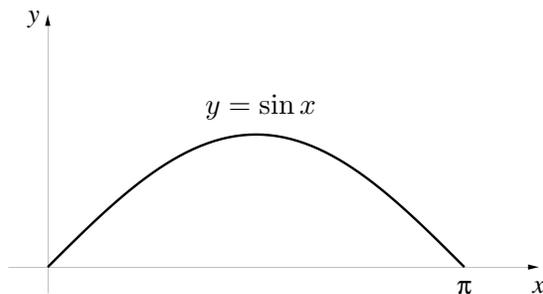
Energy used is the area under the power graph

Although we can always visualize an integral as an area, it may not be very enlightening in particular circumstances. For example, in the last section (page 358) we estimated the length of the graph $y = \sin x$ from $x = 0$ to $x = \pi$. Our estimates came from Riemann sums for the function $f(x) = \sqrt{1 + \cos^2 x}$ over the interval $[0, \pi]$. These Riemann sums converge to the integral

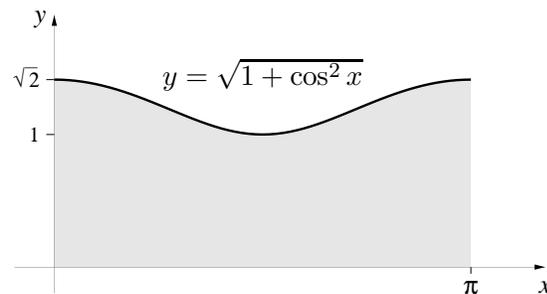
The area interpretation is not always helpful!

$$\int_0^\pi \sqrt{1 + \cos^2 x} \, dx,$$

which we can now view as the area under the graph of $\sqrt{1 + \cos^2 x}$.



length of graph = 3.82019...



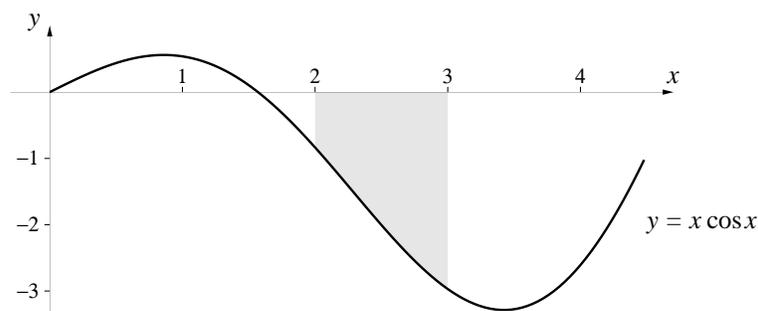
shaded area = 3.82019...

More generally, the length of $y = g(x)$ will always equal the area under $y = \sqrt{1 + (g'(x))^2}$.

The integral of a negative function

Up to this point, we have been dealing with a function $f(x)$ that is never negative on the interval $[a, b]$: $f(x) \geq 0$. Its graph therefore lies entirely

above the x -axis. What happens if $f(x)$ *does* take on negative values? We'll first consider an example.

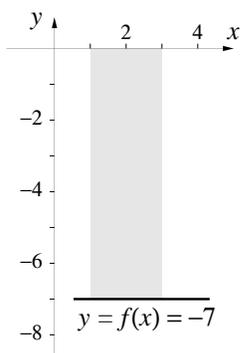


The graph of $f(x) = x \cos x$ is shown above. On the interval $[2, 3]$, it lies entirely below the x -axis. As you can check,

$$\int_2^3 x \cos x \, dx = -1.969080.$$

The integral is negative, but areas are positive. Therefore, it seems we can't interpret *this* integral as an area. But there is more to the story. The shaded region is 1 unit wide and varies in height from 1 to 3 units. If we say the average height is about 2 units, then the area is about 2 square units. Except for the negative sign, our rough estimate for the area is almost exactly the value of the integral.

In fact, the integral of a *negative* function is always the *negative* of the area between its graph and the x -axis. To see why this is always true, we'll look first at the simplest possibility—a constant function. The integral



of a constant function is just the product of that constant value by the width of the interval. For example, suppose $f(x) = -7$ on the interval $[1, 3]$. The region between the graph and the x -axis is a rectangle whose area is $7 \times 2 = 14$. However,

$$f(x) \cdot \Delta x = -7 \times 2 = -14.$$

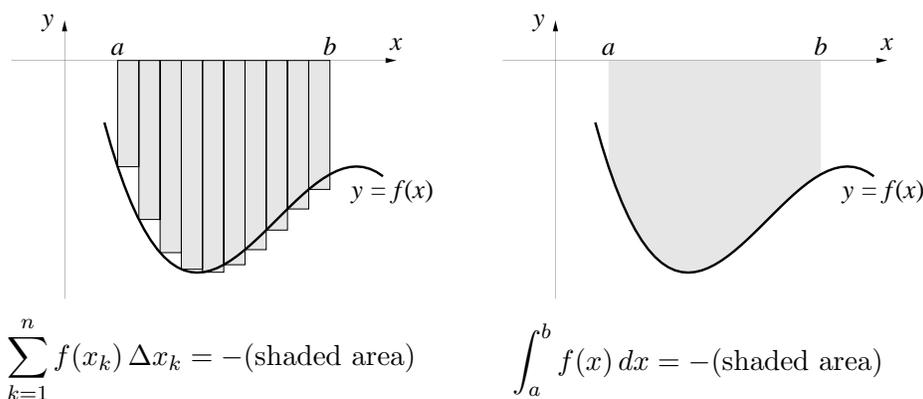
This is the *negative* of the area of the region.

Let's turn now to an arbitrary function $f(x)$ whose values vary but remain negative over the interval $[a, b]$. Each term in the Riemann sum on the left

The integral is the
negative of the area

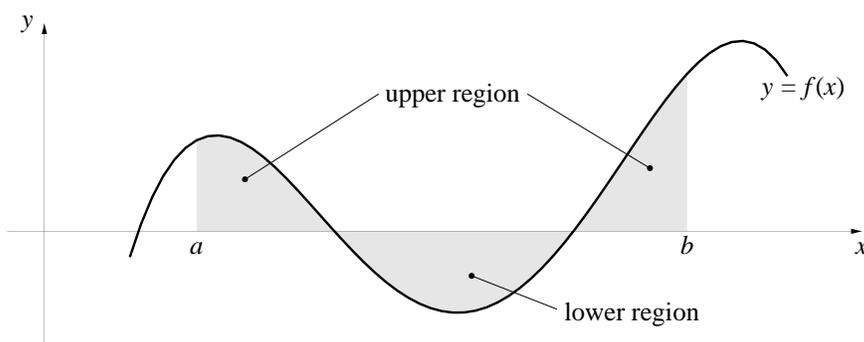
A function that takes
only negative values

is the *negative* of one of the shaded rectangles. In the process of refinement, the total area of the rectangles approaches the shaded area on the right. At the same time, the Riemann sums approach the integral. Thus, the integral must be the negative of the shaded area.



Functions with both positive and negative values

The final possibility to consider is that $f(x)$ takes both positive and negative values on the interval $[a, b]$. In that case its graph lies partly above the x -axis and partly below. By considering these two parts separately we can see that

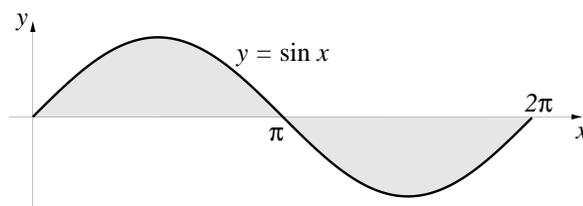


$$\int_a^b f(x) dx = (\text{area of upper region}) - (\text{area of lower region}).$$

The graph of $y = \sin x$ on the interval $[0, \pi]$ is the mirror image of the graph on the interval $[\pi, 2\pi]$. The first half lies above the x -axis, the second

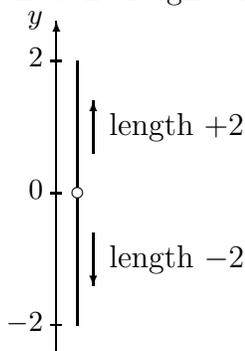
half below. Since the upper and lower areas are equal, it follows that

$$\int_0^{2\pi} \sin x \, dx = 0.$$



Signed area

There is a way to simplify the geometric interpretation of an integral as an area. It involves introducing the notion of *signed area*, by analogy with the notion of signed length.

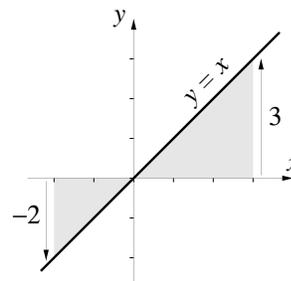


Consider the two points 2 and -2 on the y -axis at the left. Although the line that goes up from 0 to 2 has the same length as the line that goes down from 0 to -2 we customarily attach a sign to those lengths to take into account the *direction* of the line. Specifically, we assign a positive length to a line that goes up and a negative length to a line that goes down. Thus the line from 0 to -2 has **signed length** -2 .

To adapt this pattern to areas, just assign to any area that goes up from the x -axis a positive value and to any area that goes down from the x -axis a negative value. Then the **signed area** of a region that is partly above and partly below the x -axis is just the sum of the areas of the parts—taking the signs of the different parts into account.

Consider, for example, the graph of $y = x$ over the interval $[-2, 3]$. The upper region is a triangle whose area is 4.5. The lower region is another triangle; its area 2, and its *signed* area is -2 . Thus, the total signed area is $+2.5$, and it follows that

$$\int_{-2}^3 x \, dx = 2.5.$$



A region below the x -axis has negative area

You should confirm that Riemann sums for $f(x) = x$ over the interval $[-2, 3]$ converge to the value 2.5 (see exercise 13).

Now that we can describe the signed area of a region in the x, y -plane, we have a simple and uniform way to visualize the integral of any function:

$$\int_a^b f(x) dx = \text{the signed area}$$

between the graph of $f(x)$ and the x -axis.

Error Bounds

A Riemann sum determines the value of an integral only approximately. For example,

$$\int_1^3 \sqrt{1+x^3} dx = 6.229959\dots,$$

but a left endpoint Riemann sum with 100 equal subintervals Δx gives

$$\sum_{k=1}^{100} \sqrt{1+x^3} \Delta x = 6.191236.$$

If we use this sum as an estimate for the value of the integral, we make an **error** of

$$6.229959 - 6.191236 = .038723.$$

By increasing the number of subintervals, we can reduce the size of the error. For example, with 100,000 subintervals, the error is only .000038. (This information comes from pages 373–375.) The fact that the first four digits in the error are now 0 means, roughly speaking, that the first four digits in the new estimate are correct.

In this example, we could measure the error in a Riemann sum because we knew the value of the integral. Usually, though, we *don't* know the value of the integral—that's why we're calculating Riemann sums! We will describe here a method to decide how inaccurate a Riemann sum is *without first knowing the value of the integral*. For example, suppose we estimate the value of

$$\int_0^1 e^{-x^2} dx$$

The error in
a Riemann sum

Finding the error
without knowing
the exact value

Error bounds

using a left endpoint Riemann sum with 1000 equal subintervals. The value we get is .747140. Our method will tell us that this differs from the true value of the integral by *no more than* .000633. So the method does not tell us the exact size of the error. It says only that the error is not larger than .000633. Such a number is called an **error bound**. The actual error—that is, the true difference between the value of the integral and the value of the Riemann sum—may be a lot less than .000633. (That is indeed the case. In the exercises you are asked to show that the actual error is about half this number.)

We have two ways to indicate that .747140 is an estimate for the value of the integral, with an error bound of .000633. One is to use a “plus-minus” sign (\pm):

$$\int_0^1 e^{-x^2} dx = .747140 \pm .000633.$$

Since $.747140 - .000633 = .746507$ and $.747140 + .000633 = .747773$, this is the same as

$$.746507 \leq \int_0^1 e^{-x^2} dx \leq .747773.$$

Upper and lower bounds

The number .746507 is called a **lower bound** for the integral, and .747773 is called an **upper bound**. Thus, the true value of the integral is .74... , and the third digit is either a 6 or a 7.

Our method will tell us even more. In this case, it will tell us that the original Riemann sum is *already* larger than the integral. In other words, we can drop the upper bound from .747773 to .747140:

$$.746507 \leq \int_0^1 e^{-x^2} dx \leq .747140.$$

The method

Visualize the error as an area

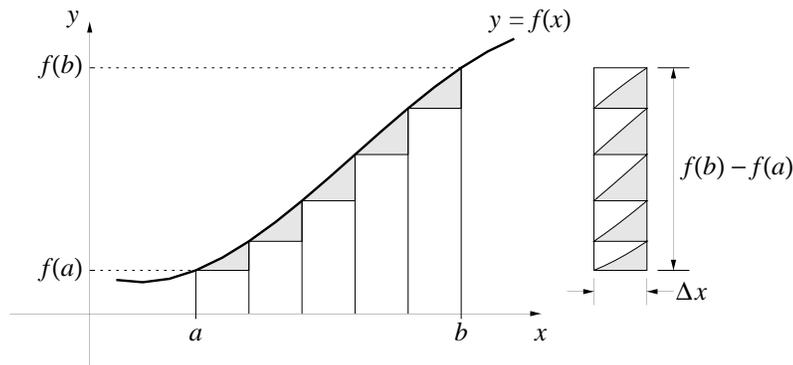
We want to get a bound on the difference between the integral of a function and a Riemann sum for that function. By visualizing both the integral and the Riemann sum as areas, we can visualize their difference as an area, too. We'll assume that all subintervals in the Riemann sum have the same width Δx . This will help keep the details simple. Thus

$$\text{error} = \left| \int_a^b f(x) dx - \sum_{k=1}^n f(x_k) \Delta x \right|.$$

(The absolute value $|u - v|$ of the difference tells us how far apart u and v are.)

We'll also assume that the function $f(x)$ is *positive* and *increasing* on an interval $[a, b]$. We say $f(x)$ is **increasing** if its graph rises as x goes from left to right.

Let's start with left endpoints for the Riemann sum. Because $f(x)$ is increasing, the rectangles lie entirely below the graph of the function:



The integral is the area under the graph, and the Riemann sum is the area of the rectangles. Therefore, the error is the area of the shaded region that lies above the rectangles and below the graph. This region consists of a number of separate pieces that sit on top of the individual rectangles. Slide them to the right and stack them on top of one another, as shown in the figure. They fit together inside a single rectangle of width Δx and height $f(b) - f(a)$. Thus the error (which is the area of the shaded region) is no greater than the area of this rectangle:

Slide the
little errors
into a stack

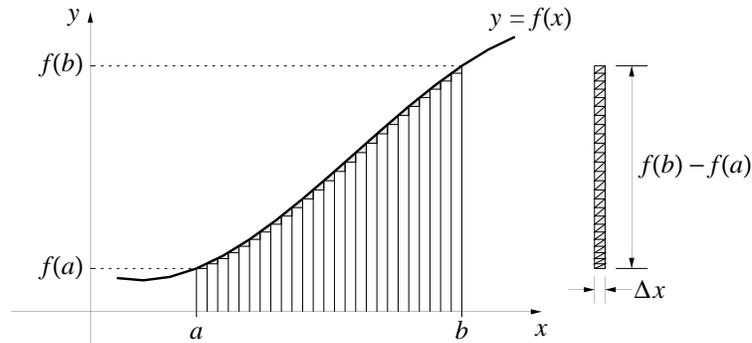
$$\text{error} \leq \Delta x (f(b) - f(a)).$$

The number $\Delta x (f(b) - f(a))$ is our error bound. It is clear from the figure that this is not the *exact* value of the error. The error is smaller, but it is difficult to say exactly how much smaller. Notice also that we need very little information to find the error bound—just Δx and the function values $f(a)$ and $f(b)$. We do *not* need to know the exact value of the integral!

The error bound is proportional to Δx . If we cut Δx in half, that will cut the error bound in half. If we make Δx a tenth of what it had been, that will make the error bound a tenth of what it had been. In the figure below, Δx is $1/5$ -th its value in the previous figure. The rectangle on the right shows how

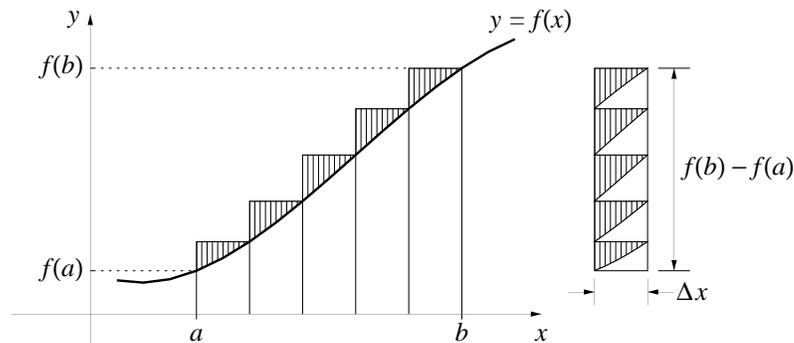
The error bound is
proportional to Δx

much smaller the error bound has become as a result. It demonstrates how Riemann sums converge as the size of the subinterval Δx shrinks to zero.



Right endpoints

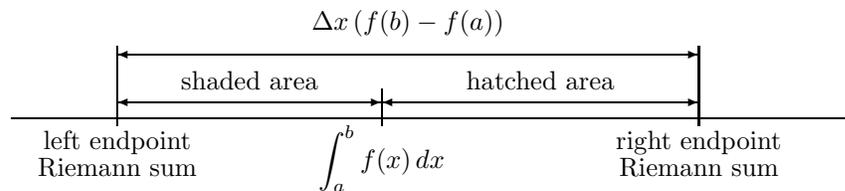
Let's go back to the original Δx , but switch to *right* endpoints. Then the tops of the rectangles lie above the graph. The error is the vertically hatched region between the graph and the tops of the rectangles. Once again, we can slide the little errors to the right and stack them on top of one another.



They fit inside the same rectangle we had before. Thus, whether the Riemann sum is constructed with left endpoints or right endpoints, we find the same error bound:

$$\text{error} \leq \Delta x (f(b) - f(a)).$$

Because $f(x)$ is increasing, the left endpoint Riemann sum is smaller than the integral, while the right endpoint Riemann sum is larger. On a number line, these three values are arranged as follows:



The distance from the left endpoint Riemann sum to the integral is represented by the shaded area, and the distance from the right endpoint by the hatched area. Notice that these two areas exactly fill the rectangle that gives us the error bound. Thus the distance between the two Riemann sums on the number line is exactly $\Delta x (f(b) - f(a))$, as shown.

Finally, suppose that the Riemann sum has *arbitrary* sampling points x_k :

Arbitrary
sampling points

$$\sum_{k=1}^n f(x_k) \Delta x.$$

Since f is increasing on the interval $[a, b]$, its values get larger as x goes from left to right. Therefore, on the k -th subinterval,

$$f(\text{left endpoint}) \leq f(x_k) \leq f(\text{right endpoint}).$$

In other words, the rectangle built over the left endpoint is the shortest, and the one built over the right endpoint is the tallest. The one built over the sampling point x_k lies somewhere in between.

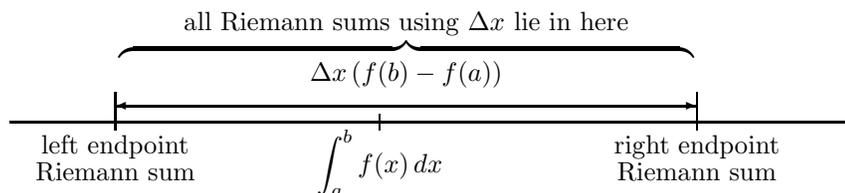
The areas of these three rectangles are arranged in the same order:

$$f(\text{left endpoint}) \Delta x \leq f(x_k) \Delta x \leq f(\text{right endpoint}) \Delta x.$$

If we add up these areas, we get Riemann sums. The Riemann sums are arranged in the same order as their individual terms:

$$\left\{ \begin{array}{l} \text{left endpoint} \\ \text{Riemann sum} \end{array} \right\} \leq \sum_{k=1}^n f(x_k) \Delta x \leq \left\{ \begin{array}{l} \text{right endpoint} \\ \text{Riemann sum} \end{array} \right\}.$$

Thus, the left endpoint and the right endpoint Riemann sums are extremes: *every* Riemann sum for f that uses a subinterval size of Δx lies between these two.



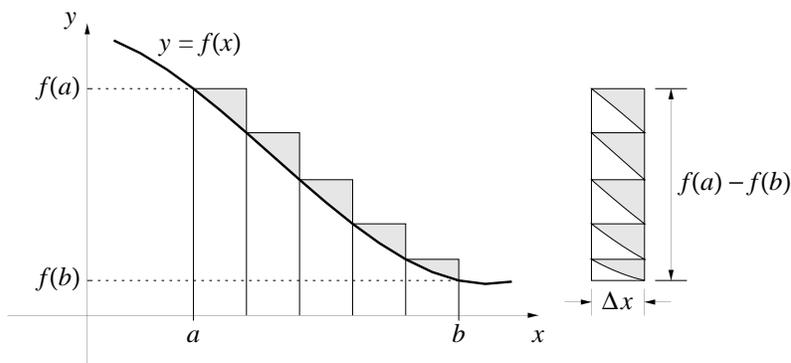
It follows that $\Delta x (f(b) - f(a))$ is an error bound for all Riemann sums whose subintervals are Δx units wide.

Error bounds for a decreasing function

If $f(x)$ is a positive function but is *decreasing* on the interval $[a, b]$, we get essentially the same result. Any Riemann sum for f that uses a subinterval of size Δx differs from the integral by no more than $\Delta x (f(a) - f(b))$. When f is decreasing, however, $f(a)$ is larger than $f(b)$, so we write the height of the rectangle as $f(a) - f(b)$. To avoid having to pay attention to this distinction, we can use absolute values to describe the error bound:

$$\text{error} \leq \Delta x |f(b) - f(a)|.$$

Furthermore, when f is decreasing, the left endpoint Riemann sum is larger than the integral, while the right endpoint Riemann sum is smaller. The difference between the right and the left endpoint Riemann sums is still $\Delta x |f(b) - f(a)|$.

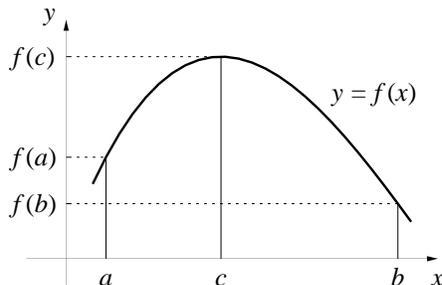


Monotonic functions

Up to this point, we have assumed $f(x)$ was either always increasing, or else always decreasing, on the interval $[a, b]$. Such a function is said to be **monotonic**. If $f(x)$ is *not* monotonic, the process of getting an error bound for Riemann sums is only slightly more complicated.

Error bounds for non-monotonic functions

Here is how to get an error bound for the Riemann sums constructed for a non-monotonic function. First break up the interval $[a, b]$ into smaller pieces on which the function *is* monotonic.



In this figure, there are two such intervals: $[a, c]$ and $[c, b]$. Suppose we construct a Riemann sum for f by using rectangles of width Δx_1 on the first interval, and Δx_2 on the second. Then the total error for this sum will be no larger than the sum of the error bounds on the two intervals:

The monotonic pieces

$$\text{total error} \leq \Delta x_1 |f(c) - f(a)| + \Delta x_2 |f(b) - f(c)|.$$

By making Δx_1 and Δx_2 sufficiently small, we can make the error as small as we wish.

This method can be applied to any non-monotonic function that can be broken up into monotonic pieces. For other functions, more than two pieces may be needed.

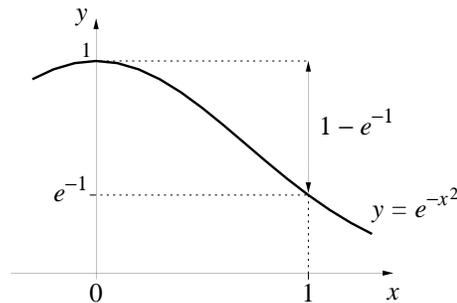
Using the method

Earlier we said that when we use a left endpoint Riemann sum with 1000 equal subintervals to estimate the value of the integral

An example

$$\int_0^1 e^{-x^2} dx,$$

the error is no larger than .000633. Let's see how our method would lead to this conclusion.



On the interval $[0, 1]$, $f(x) = e^{-x^2}$ is a decreasing function. Furthermore,

$$f(0) = 1 \quad f(1) = e^{-1} \approx .3679.$$

If we divide $[0, 1]$ into 1000 equal subintervals Δx , then $\Delta x = 1/1000 = .001$. The error bound is therefore

$$.001 \times |.3679 - 1| = .001 \times .6321 = .0006321.$$

Any number *larger* than this one will also be an error bound. By “rounding up,” we get .000633. This is slightly shorter to write, and it is the bound we claimed earlier. An even shorter bound is .00064.

Furthermore, since $f(x) = e^{-x^2}$ is decreasing, any Riemann sum constructed with *left* endpoints is larger than the actual value of the integral. Since the left endpoint Riemann sum with 1000 equal subdivisions has the value .747140, upper and lower bounds for the integral are

$$.746507 = .747140 - .000633 \leq \int_0^1 e^{-x^2} dx \leq .747140.$$

Integration Rules

Just as there are rules that tell us how to find the derivative of various combinations of functions, there are other rules that tell us how to find the integral. Here are three that are exactly analogous to differentiation rules.

$$\begin{aligned} \int_a^b f(x) + g(x) dx &= \int_a^b f(x) dx + \int_a^b g(x) dx \\ \int_a^b f(x) - g(x) dx &= \int_a^b f(x) dx - \int_a^b g(x) dx \\ \int_a^b c f(x) dx &= c \int_a^b f(x) dx \end{aligned}$$

Compare
Riemann sums

Let’s see why the third rule is true. A Riemann sum for the integral on the left looks like

$$\sum_{k=1}^n c f(x_k) \Delta x_k = c f(x_1) \Delta x_1 + \cdots + c f(x_n) \Delta x_n.$$

Since the factor c appears in every term in the sum, we can move it outside the summation:

$$c(f(x_1) \Delta x_1 + \cdots + f(x_n) \Delta x_n) = c \left(\sum_{k=1}^n f(x_k) \Delta x_k \right).$$

The new expression is c times a Riemann sum for

$$\int_a^b f(x) dx.$$

Since the Riemann sum expressions are equal, the integral expressions they converge to must be equal, as well. You can use similar arguments to show why the other two rules are true.

Here is one example of the way we can use these rules:

Using the rules

$$\int_1^3 4\sqrt{1+x^3} dx = 4 \int_1^3 \sqrt{1+x^3} dx = 4 \times 6.229959 = 24.919836.$$

(The value of the second integral is given on page 385.) Here is another example:

$$\int_2^9 5x^7 - 2x^3 + 24x dx = 5 \int_2^9 x^7 dx - 2 \int_2^9 x^3 dx + 24 \int_2^9 x dx.$$

Of course, we must still determine the value of various integrals of the form

$$\int_a^b x^n dx.$$

However, the example shows us that, once we know the value of these special integrals, we can determine the value of the integral of any polynomial.

Here are two more rules that have no direct analogue in differentiation. The first says that if $f(x) \leq g(x)$ for every x in the interval $[a, b]$, then

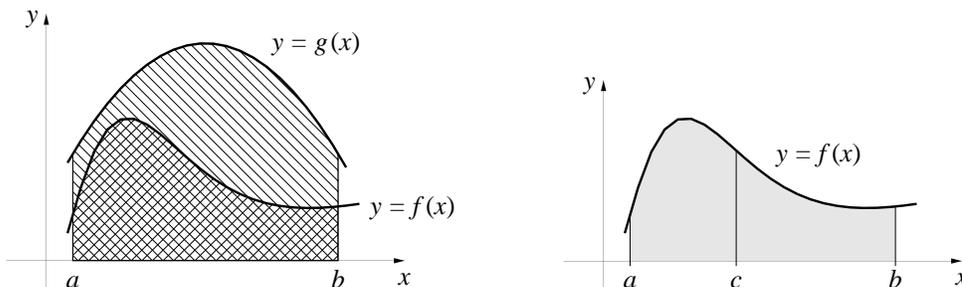
Two more rules

$$\int_a^b f(x) dx \leq \int_a^b g(x) dx.$$

In the second, c is a point somewhere in the interval $[a, b]$:

$$\int_a^b f(x) dx = \int_a^c f(x) dx + \int_c^b f(x) dx.$$

If you visualize an integral as an area, it is clear why these rules are true.



The second rule can be used to understand the results of exercise 4 on page 366. It concerns the three integrals

$$\int_1^3 \sqrt{1+x^3} dx = 6.229959, \quad \int_3^7 \sqrt{1+x^3} dx = 45.820012,$$

and

$$\int_1^7 \sqrt{1+x^3} dx = 52.049971.$$

Since the third interval $[1, 7]$ is just the first $[1, 3]$ combined with the second $[3, 7]$, the third integral is the sum of the first and the second. The numerical values confirm this.

Exercises

1. Determine the values of the following integrals.

$$\begin{array}{lll} \text{a) } \int_2^{15} 3 dx & \text{c) } \int_{-5}^{-3} 7 dx & \text{e) } \int_{-4}^9 -2 dz \\ \text{b) } \int_{15}^2 3 dx & \text{d) } \int_{-3}^{-5} 7 dx & \text{f) } \int_9^{-4} -2 dz \end{array}$$

2. a) Sketch the graph of

$$g(x) = \begin{cases} 7 & \text{if } 1 \leq x < 5, \\ -3 & \text{if } 5 \leq x \leq 10 \end{cases}$$

b) Determine $\int_1^7 g(x) dx$, $\int_7^{10} g(x) dx$, and $\int_1^{10} g(x) dx$.

Refining Riemann sums

3. a) By refining Riemann sums, find the value of the following integral to four decimal places accuracy. Do the computations twice: first, using left endpoints; second, using midpoints.

$$\int_0^1 \frac{1}{1+x^3} dx.$$

b) How many subintervals did you need to get four decimal places accuracy when you used left endpoints and when you used midpoints? Which sampling points gave more efficient computations—left endpoints or midpoints?

4. By refining appropriate Riemann sums, determine the value of each of the following integrals, accurate to four decimal places. Use whatever sampling points you wish, but justify your claim that your answer is accurate to four decimal places.

$$\text{a) } \int_1^4 \sqrt{1+x^3} dx \qquad \text{b) } \int_4^7 \sqrt{1+x^3} dx \qquad \text{c) } \int_0^3 \frac{\cos x}{1+x^2} dx$$

[Answer: $\int_0^3 \frac{\cos x}{1+x^2} dx = .6244\dots$]

5. Determine the value of the following integrals to four decimal places accuracy.

$$\begin{array}{ll} \text{a) } \int_1^2 e^{-x^2} dx & \text{c) } \int_0^4 \sin(x^2) dx \\ \text{b) } \int_0^4 \cos(x^2) dx & \text{d) } \int_0^1 \frac{4}{1+x^2} dx \end{array}$$

6. a) What is the length of the graph of $y = \sqrt{x}$ from $x = 1$ to $x = 4$?

b) What is the length of the graph of $y = x^2$ from $x = 1$ to $x = 2$?

c) Why are the answers in parts (a) and (b) the same?

7. Both of the curves $y = 2^x$ and $y = 1 + x^{3/2}$ pass through $(0, 1)$ and $(1, 2)$. Which is the shorter one? Can you decide simply by looking at the graphs?

8. A pyramid is 30 feet tall. The area of a horizontal cross-section x feet from the top of the pyramid measures $2x^2$ square feet. What is the area of the base? What is the volume of the pyramid, to the nearest cubic foot?

Error bounds

9. A left endpoint Riemann sum with 1000 equally spaced subintervals gives the estimate .135432 for the value of the integral

$$\int_1^2 e^{-x^2} dx.$$

- a) Is the true value of the integral larger or smaller than this estimate? Explain.
- b) Find an error bound for this estimate.
- c) Using the information you have already assembled, find lower and upper bounds A and B :

$$A \leq \int_1^2 e^{-x^2} dx \leq B.$$

- d) The lower and upper bounds allow you to determine a certain number of digits in the *exact* value of the integral. How many digits do you know, and what are they?

10. A left endpoint Riemann sum with 100 equally spaced subintervals gives the estimate .342652 for the value of the integral

$$\int_0^{\pi/4} \tan x dx.$$

- a) Is the true value of the integral larger or smaller than this estimate? Explain.
- b) Find an error bound for this estimate.
- c) Using the information you have already assembled, find lower and upper bounds A and B :

$$A \leq \int_0^{\pi/4} \tan x dx \leq B.$$

- d) The lower and upper bounds allow you to determine a certain number of digits in the *exact* value of the integral. How many digits do you know, and what are they?

11. a) In the next section you will see that

$$\int_0^{\pi/2} \sin x dx = 1$$

exactly. Here, estimate the value by a Riemann sum using the left endpoints of 100 equal subintervals.

b) Find an error bound for this estimate, and use it to construct the best possible lower and upper bounds

$$A \leq \int_0^{\pi/2} \sin x \, dx \leq B.$$

12. In the text (page 373), a Riemann sum using left endpoints on 1000 equal subintervals produces an estimate of 6.226083 for the value of

$$\int_1^3 \sqrt{1+x^3} \, dx.$$

a) Is the true value of the integral larger or smaller than this estimate? Explain your answer, and do so without referring to the fact that the true value of the integral is known to be 6.229959...

b) Find an error bound for this estimate.

c) Find the upper and lower bounds for the value of the integral that are determined by this estimate.

d) According to these bounds, how many digits of the value of the integral are now known for certain?

The average value of a function

In the exercises for section 1 we saw that the average staffing level for a job is

$$\text{average staffing} = \frac{\text{total staff-hours}}{\text{hours worked}}.$$

If $S(t)$ represents the number of staff working at time t , then the total staff-hours accumulated between $t = a$ and $t = b$ hours is

$$\text{total staff-hours} = \int_a^b S(t) \, dt \quad \text{staff-hours.}$$

Therefore the average staffing is

$$\text{average staffing} = \frac{1}{b-a} \int_a^b S(t) \, dt \quad \text{staff.}$$

Likewise, if a town's power demand was $p(t)$ megawatts at t hours, then its average power demand between $t = a$ and $t = b$ hours is

$$\text{average power demand} = \frac{1}{b-a} \int_a^b p(t) dt \quad \text{megawatts.}$$

Geometric meaning of the average

We can define the **average value** of an arbitrary function $f(x)$ over an interval $a \leq x \leq b$ by following this pattern:

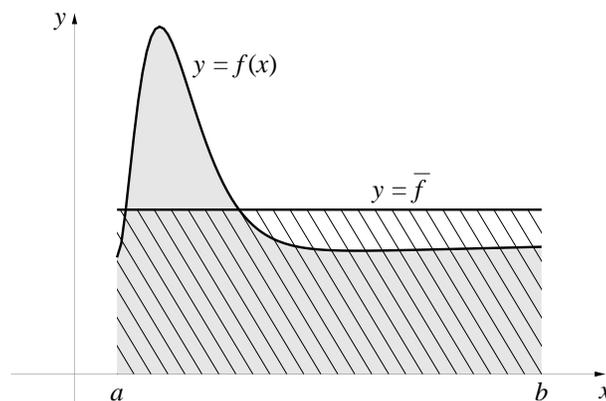
$$\text{average value of } f = \frac{1}{b-a} \int_a^b f(x) dx.$$

The average value of f is sometimes denoted \bar{f} . Since

$$\bar{f} \cdot (b-a) = \int_a^b f(x) dx,$$

the area under the horizontal line $y = \bar{f}$ between $x = a$ and $x = b$ is the same as the area under the graph $y = f(x)$. See the graph below.

13. a) What is the average value of $f(x) = 5$ on the interval $[1, 7]$?
 b) What is the average value of $f(x) = \sin x$ on the interval $[0, 2\pi]$? On $[0, 100\pi]$?
 c) What is the average value of $f(x) = \sin^2 x$ on the interval $[0, 2\pi]$? On $[0, 100\pi]$?



14. a) What is the average value of the function

$$H(x) = \begin{cases} 1 & \text{if } 0 \leq x < 4, \\ 12 & \text{if } 4 \leq x < 6, \\ 1 & \text{if } 6 \leq x \leq 20, \end{cases}$$

on the interval $[0, 20]$?

b) Is the average \overline{H} larger or smaller than the average of the two numbers 12 and 1 that represent the largest and smallest values of the function?

c) Sketch the graph $y = H(x)$ along with the horizontal line $y = \overline{H}$, and show directly that the same area lies under each of these two graphs over the interval $[0, 20]$.

15. a) What are the maximum and minimum values of $f(x) = x^2e^{-x}$ on the interval $[0, 20]$? What is the *average* of the maximum and the minimum?

b) What is the average \overline{f} of $f(x)$ on the interval $[0, 20]$?

c) Why aren't these two averages the same?

The integral as a signed area

16. By refining Riemann sums, confirm that $\int_{-2}^3 x \, dx = 2.5$.

17. a) Sketch the graphs of $y = \cos x$ and $y = 5 + \cos x$ over the interval $[0, 4\pi]$.

b) Find $\int_0^{4\pi} \cos x \, dx$ by visualizing the integral as a signed area.

c) Find $\int_0^{4\pi} 5 + \cos x \, dx$. Why does $\int_0^{4\pi} 5 \, dx$ have the same value?

18. a) By refining appropriate Riemann sums, determine the value of the integral $\int_0^{\pi} \sin^2 x \, dx$ to four decimal places accuracy.

b) Sketch the graph of $y = \sin^2 x$ on the interval $0 \leq x \leq \pi$. Note that your graph lies inside the rectangle formed by the lines $y = 0$, $y = 1$, $x = 0$ and $x = \pi$. (Sketch this rectangle.)

c) Explain why the area under the graph of $y = \sin^2 x$ is exactly *half* of the area of the rectangle you sketched in part (b). What is the area of that rectangle?

d) Using your observations in part (c), explain why $\int_0^\pi \sin^2 x \, dx$ is *exactly* $\pi/2$.

19. a) On what interval $a < x < b$ does the graph of the function $y = 4 - x^2$ lie *above* the x -axis?

b) Sketch the graph of $y = 4 - x^2$ on the interval $a \leq x \leq b$ you determined in part (a).

c) What is the area of the region that lies above the x -axis and below the graph of $y = 4 - x^2$?

20. a) What is the signed area (see page 384 in the text) between the graph of $y = x^3 - x$ and the x -axis on the interval $-1 \leq x \leq 2$?

b) Sketch the graph of $y = x^3 - x$ on the interval $-1 \leq x \leq 2$. On the basis of your sketch, support or refute the following claim: the *signed* area between the graph of $y = x^3 - x$ and the x -axis on the interval $-1 \leq x \leq 2$ is exactly the same as the area between the graph of $y = x^3 - x$ and the x -axis on the interval $+1 \leq x \leq 2$.

6.4 The Fundamental Theorem of Calculus

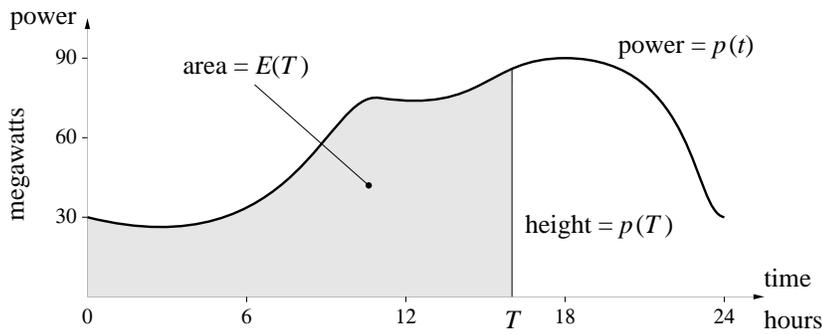
Two Views of Power and Energy

In section 1 we considered how much energy a town consumed over 24 hours when it was using $p(t)$ megawatts of power at time t . Suppose $E(T)$ megawatt-hours of energy were consumed during the first T hours. Then the integral, introduced in section 3, gave us the language to describe how E depends on p :

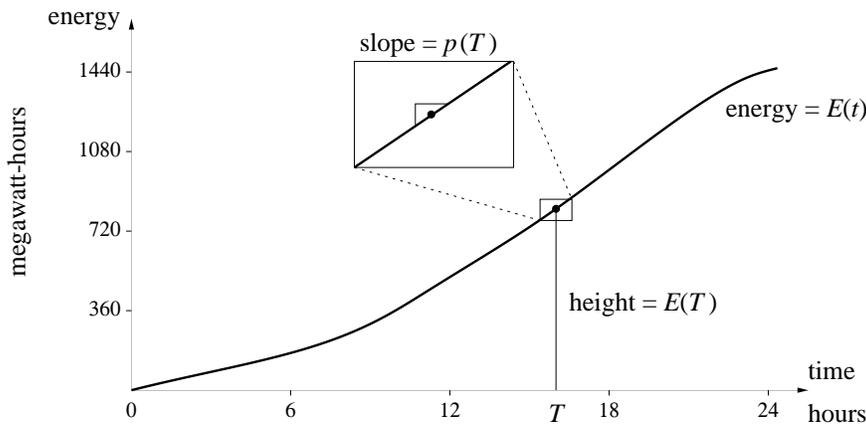
Energy is the integral of power...

$$E(T) = \int_0^T p(t) dt.$$

Because E is the integral of p , we can visualize $E(T)$ as the area under the power graph $y = p(t)$ as the time t sweeps from 0 hours to T hours:



As T increases, so does $E(T)$. The exact relation between E and T is shown in the graph below. The height of the graph of E at any point T is equal to the area under the graph of p from 0 out to the point T .



The microscope window on the graph of E reminds us that the slope of the graph at any point T is just $p(T)$:

$$p(T) = E'(T).$$

... and power is the derivative of energy

We discovered this fact in section 1, where we stated it in the following form: *power is the rate at which energy is consumed.*

Pause now to study the graphs of power and energy. You should convince yourself that the *height* of the E graph at any time equals the *area* under the p graph up to that time. For example, when $T = 0$ no area has accumulated, so $E(0) = 0$. Furthermore, up to $T = 6$ hours, power demand was almost constant at about 30 megawatts. Therefore, $E(6)$ should be about

$$30 \text{ megawatts} \times 6 \text{ hours} = 180 \text{ megawatt-hours.}$$

It is. You should also convince yourself that the *slope* of the E graph at any point equals the *height* of the p graph at that point. Thus, for example, the graph of E will be steepest where the graph of p is tallest.

Notice that when we write the energy accumulation function $E(T)$ as an integral,

$$E(T) = \int_0^T p(t) dt,$$

A new ingredient: variable limits of integration

we have introduced a new ingredient. The time variable T appears as one of the “limits of integration.” By definition, the integral is a single number. However, that number depends on the interval of integration $[0, T]$. As soon as we treat T as a variable, the integral itself becomes a variable, too. Here the value of the integral varies with T .

We now have two ways of viewing the relation between power and energy. According to the first view, the energy accumulation function E is the integral of power demand:

$$E(T) = \int_0^T p(t) dt.$$

This integral is a solution to a differential equation

According to the second view, the energy accumulation function is the solution $y = E(t)$ to an initial value problem defined by power demand:

$$y' = p(t); \quad y(0) = 0.$$

If we take the first view, then we find E by refining Riemann sums—because that is the way to determine the value of an integral. If we take the second

view, then we can find E by using any of the methods for solving initial value problems that we studied in chapter 4. Thus, the energy integral is a solution to a certain differential equation.

This is unexpected. Differential equations involve derivatives. At first glance, they have nothing to do with integrals. Nevertheless, the relation between power and energy shows us that there is a deep connection between derivatives and integrals. As we shall see, the connection holds for the integral of *any* function. The connection is so important—because it links together the two basic processes of calculus—that it has been called the **fundamental theorem of calculus**.

The fundamental theorem of calculus

The fundamental theorem gives us a powerful new tool to calculate integrals. Our aim in this section is to see why the theorem is true, and to begin to explore its use as a tool. In Calculus II we will consider many specific integration techniques that are based on the fundamental theorem.

Integrals and Differential Equations

We begin with a statement of the fundamental theorem for a typical function $f(x)$.

The Fundamental Theorem of Calculus

The solution $y = A(x)$ to the initial value problem

$$y' = f(x) \quad y(a) = 0$$

is the **accumulation function** $A(X) = \int_a^X f(x) dx$.

We have always been able to find the value of the integral

$$\int_a^b f(x) dx$$

by refining Riemann sums. The fundamental theorem gives us a new way. It says: First, find the solution $y = A(x)$ to the initial value problem $y' = f(x)$, $y(a) = 0$ using any suitable method for solving the differential equation. Then, once we have $A(x)$, we get the value of the integral by evaluating A at $x = b$.

A test case

To see how all this works, let's find the value of the integral

$$\int_0^4 \cos(x^2) dx$$

two ways: by refining Riemann sums and by solving the initial value problem

$$y' = \cos(x^2); \quad y(0) = 0$$

using Euler's method.

RIEMANN versus...

To estimate the value of the integral, we use the program RIEMANN from page 362. It produces the sequence of left endpoint sums shown in the table below on the left. Since the first three digits have stabilized, the value of the integral is .594... . The integral was deliberately chosen to lead to the initial value problem we first considered on page 267. We solved that problem by Euler's method, using the program TABLE, and produced the table of estimates for the value of $y(4)$ that appear in the table on the right. We see $y(4) = .594\dots$

...TABLE

Program: RIEMANN
Left endpoint Riemann sums

```
DEF fnf (x) = COS(x ^ 2)
a = 0
b = 4
numberofsteps = 2 ^ 3
deltax = (b - a) / numberofsteps
x = a
accumulation = 0
FOR k = 1 TO numberofsteps
  deltaS = fnf(x) * deltax
  accumulation = accumulation + deltaS
  x = x + deltax
NEXT k
PRINT accumulation
```

Program: TABLE
Euler's method

```
DEF fnf (t) = COS(t ^ 2)
tinitial = 0
tfinal = 4
numberofsteps = 2 ^ 3
deltat = (tfinal - tinitial) / numberofsteps
t = tinitial
accumulation = 0
FOR k = 1 TO numberofsteps
  deltat = fnf(t) * deltat
  accumulation = accumulation + deltat
  t = t + deltat
NEXT k
PRINT accumulation
```

number of steps	estimated value of the integral
2^3	1.13304
2^6	.65639
2^9	.60212
2^{12}	.59542
2^{15}	.59458
2^{18}	.59448

number of steps	estimated value of $y(4)$
2^3	1.13304
2^6	.65639
2^9	.60212
2^{12}	.59542
2^{15}	.59458
2^{18}	.59448

RIEMANN estimates the value of an integral by calculating Riemann sums. TABLE solves a differential equation by constructing estimates using Euler's method. These appear to be quite different tasks, but they lead to exactly the same results! But this is no accident. Compare the programs. (In both, the PRINT statement has been put outside the loop. This speeds up the calculations but still gives us the final outcome.) Once you make necessary modifications (e.g., change `x` to `t`, `a` to `tinitial`, *et cetera*), you can see the two programs are the same.

The two programs
are the same

The very fact that these two programs do the same thing gives us one proof of the fundamental theorem of calculus.

Graphing accumulation functions

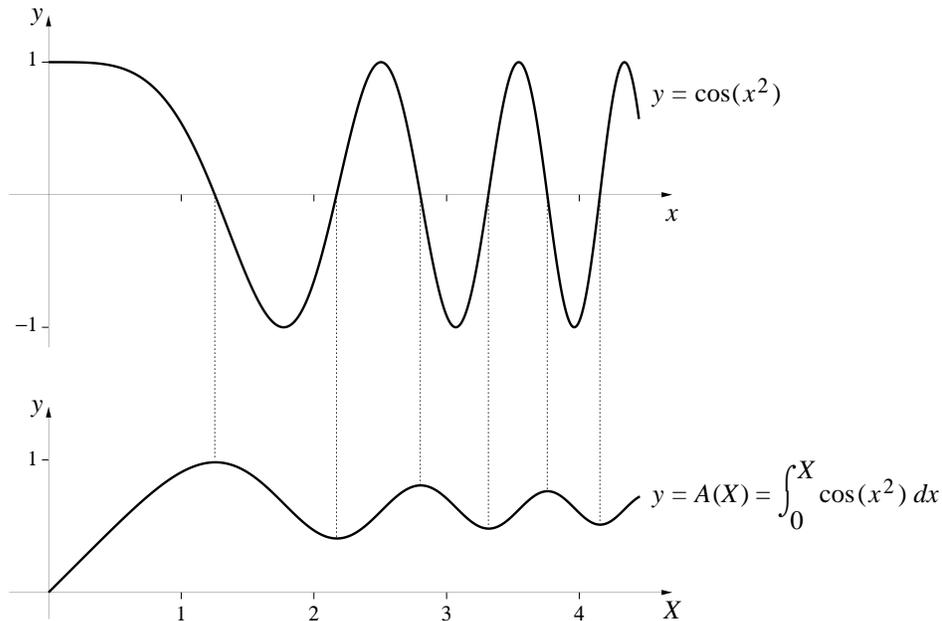
Let's take a closer look at the solution $y = A(x)$ to the initial value problem

$$y' = \cos(x^2); \quad y(0) = 0.$$

We can write it as the accumulation function

$$A(X) = \int_0^X \cos(x^2) dx.$$

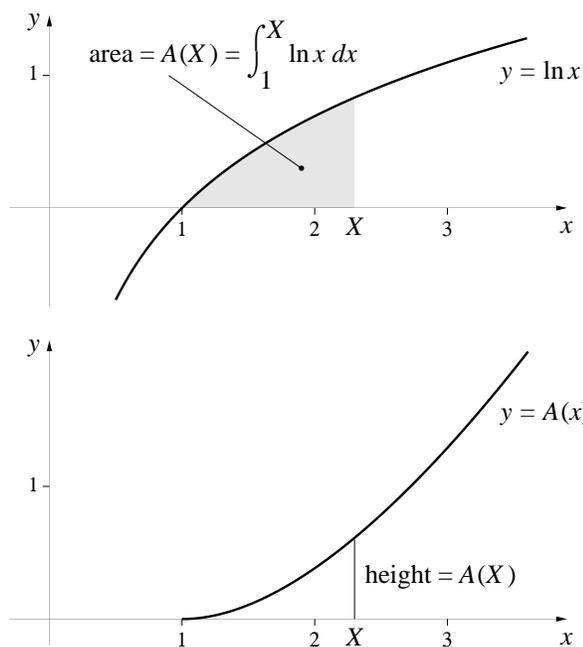
By switching to the graphing version of TABLE (this is the program PLOT—see page 269), we can graph A . Here is the result.



The relation between $y = \cos(x^2)$ and $y = A(x)$ is the same as the relation between power and energy.

- The *height* of the graph $y = A(x)$ at any point $x = X$ is equal to the *signed area* between the graph $y = \cos(x^2)$ and the x -axis over the interval $0 \leq x \leq X$.
- On the intervals where $\cos(x^2)$ is positive, $A(x)$ is increasing. On the intervals where $\cos(x^2)$ is negative, $A(x)$ is decreasing.
- When $\cos(x^2) = 0$, $A(x)$ has a maximum or a minimum.
- The *slope* of the graph of $y = A(x)$ at any point $x = X$ is equal to the *height* of the graph $y = \cos(x^2)$ at that point.

In summary, the lower curve is the integral of the upper one, and the upper curve is the derivative of the lower one.



To get a better idea of the simplicity and power of this approach to integrals, let's look at another example:

$$A(X) = \int_1^X \ln x dx.$$

The function A is the solution to the initial value problem

$$y' = \ln x, \quad y(1) = 0;$$

we find it, as usual, by Euler's method. The graphs $y = \ln x$ and $y = A(x)$ are shown at the left. Notice once again that the height of the graph of $A(x)$ at any point $x = X$ equals the area under the graph of $y = \ln x$ from $x = 1$ to $x = X$. Also, the graph of A becomes steeper as the height of $\ln x$ increases. In particular, the graph of A is horizontal (at $x = 1$) when $\ln x = 0$.

Antiderivatives

We have developed a novel approach to integration in this section. We start by replacing a given integral by an accumulation function:

Find an integral...

$$\int_a^b f(x) dx \quad \rightsquigarrow \quad A(X) = \int_a^X f(x) dx.$$

Then we try to find $A(X)$. If we do, then the original integral is just the value of A at $X = b$.

At first glance, this doesn't seem to be a sensible approach. We appear to be making the problem harder: instead of searching for a single number, we must now find an entire function. However, we know that $y = A(x)$ solves the initial value problem

$$y' = f(x), \quad y(a) = 0.$$

This means we can use the complete 'bag of tools' we have for solving differential equations to find A . The real advantage of the new approach is that it reduces integration to the fundamental activity of calculus—solving differential equations.

... by solving a differential equation ...

The differential equation $y' = f(x)$ that arises in integration problems is special. The right hand side depends only on the input variable x . We studied this differential equation in chapter 4.5, where we developed a special method to solve it—**antidifferentiation**.

... using antidifferentiation

We say that $F(x)$ is an **antiderivative** of $f(x)$ if f is the derivative of F : $F'(x) = f(x)$. Here are some examples (from page 264):

function	antiderivative
$5x^4 - 2x^3$	$x^5 - \frac{1}{2}x^4$
$5x^4 - 2x^3 + 17x$	$x^5 - \frac{1}{2}x^4 + \frac{17}{2}x^2$
$6 \cdot 10^z + 17/z^5$	$6 \cdot 10^z / \ln 10 - 17/4z^4$
$3 \sin t - 2t^3$	$-3 \cos t - \frac{1}{2}t^4$
$\pi \cos x + \pi^2$	$\pi \sin x + \pi^2 x$

Since $y = A(x)$ solves the differential equation $y' = f(x)$, we have $A'(x) = f(x)$. Thus, $A(x)$ is an antiderivative of $f(x)$, so we can try to find A by antidifferentiating f .

A is an antiderivative of f

Here is an example. Suppose we want to find

$$A(X) = \int_2^X 5x^4 - 2x^3 dx.$$

Notice that $f(x) = 5x^4 - 2x^3$ is the first function in the previous table. Since A must be an antiderivative of f , let's try the antiderivative for f that we find in the table:

$$F(x) = x^5 - \frac{1}{2}x^4.$$

Check the condition
 $A(2) = 0$

The problem is that F must also satisfy the initial condition $F(2) = 0$. However,

$$F(2) = 2^5 - \frac{1}{2} \cdot 2^4 = 32 - \frac{1}{2} \cdot 16 = 24 \neq 0,$$

so the initial condition does not hold *for this particular choice* of antiderivative. But this problem is easy to fix. Let

$$A(x) = F(x) - 24.$$

Since $A(x)$ differs from $F(x)$ only by a constant, it has the same derivative—namely, $5x^4 - 2x^3$. So $A(x)$ is still an antiderivative of $5x^4 - 2x^3$. But it also satisfies the initial condition:

$$A(2) = F(2) - 24 = 24 - 24 = 0.$$

Therefore $A(x)$ solves the problem; it has the right derivative *and* the right value at $x = 2$. Thus we have a formula for the accumulation function

$$\int_2^X 5x^4 - 2x^3 dx = X^5 - \frac{1}{2}X^4 - 24.$$

The key step in finding the correct accumulation function A was to recognize that a given function f has infinitely many antiderivatives: if F is an antiderivative, then so is $F + C$, for any constant C . The general procedure for finding an accumulation function involves these two steps:

- To find $A(X) = \int_a^X f(x) dx$:
1. first find *an* antiderivative $F(x)$ of $f(x)$,
 2. then set $A(x) = F(x) - F(a)$.

Comment: Recall that some functions f simply cannot be integrated—the Riemann sums they define may not converge. Although we have not stated it explicitly, you should keep in mind that the procedure just described applies only to functions that can be integrated.

A caution

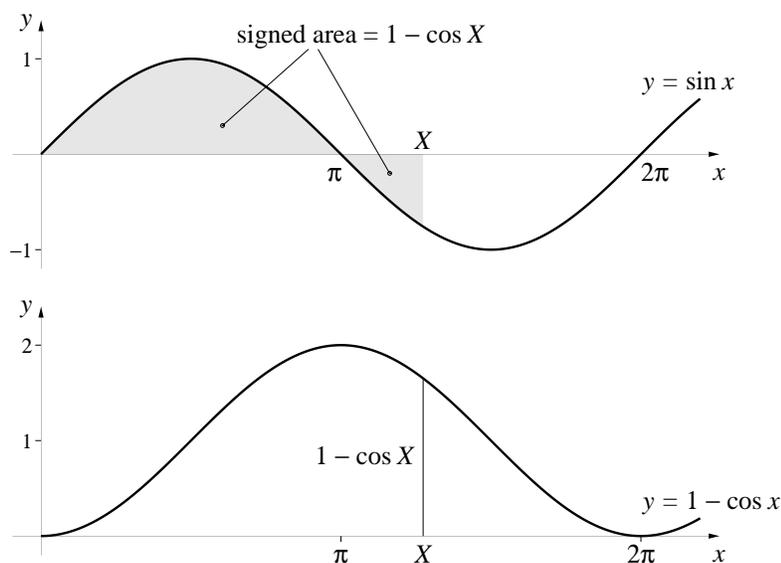
Example. To illustrate the procedure, let's find the accumulation function

$$A(X) = \int_0^X \sin x \, dx.$$

The first step is to find an antiderivative for $f(x) = \sin x$. A natural choice is $F(x) = -\cos x$. To carry out the second step, note that $a = 0$. Since $F(0) = -\cos 0 = -1$, we set

$$A(x) = F(x) - F(0) = -\cos x - (-1) = 1 - \cos x.$$

The graphs $y = \sin x$ and $y = 1 - \cos x$ are shown below.



Now that we have a formula for $A(x)$ we can find the exact value of integrals involving $\sin x$. For instance,

$$\int_0^{\pi} \sin x \, dx = A(\pi) = 1 - \cos \pi = 1 - (-1) = 2.$$

Also,

$$\int_0^{\pi/2} \sin x \, dx = A(\pi/2) = 1 - \cos \pi/2 = 1 - (0) = 1,$$

Evaluating
specific integrals

and

$$\int_0^{2\pi} \sin x \, dx = A(2\pi) = 1 - \cos 2\pi = 1 - (1) = 0.$$

These values are *exact*, and we got them without calculating Riemann sums. However, we have already found the value of the third integral. On page 383 we argued from the shape of the graph $y = \sin x$ that the signed area between $x = 0$ and $x = 2\pi$ must be 0. That argument is no help in finding the area to $x = \pi$ or to $x = \pi/2$, though. Before the fundamental theorem showed us we could evaluate integrals by finding antiderivatives, we could only make estimates using Riemann sums. Here, for example, are midpoint Riemann sums for

	subintervals	Riemann sum
$\int_0^{\pi/2} \sin x \, dx :$	10	1.001 028 868
	100	1.000 010 325
	1000	1.000 000 147
	10000	1.000 000 045

According to the table, the value of the integral is 1.000000..., to six decimal places accuracy. That is valuable information, and is often as accurate as we need. However, the fundamental theorem tells us that the value of the integral is 1 *exactly*! With the new approach, we can achieve absolute precision.

The fundamental theorem makes absolute precision possible

Precision is the result of having a *formula* for the antiderivative. Notice how we used that formula to express the value of the integral. Starting with an arbitrary antiderivative $F(x)$ of $f(x)$, we get

$$A(X) = \int_a^X f(x) \, dx = F(X) - F(a).$$

If we set $X = b$ we find

$$\int_a^b f(x) \, dx = F(b) - F(a).$$

In other words,

If $F(x)$ is any antiderivative of $f(x)$,
 then $\int_a^b f(x) \, dx = F(b) - F(a)$.

Example. Evaluate $\int_0^{\ln 2} e^{-u} du$.

For the antiderivative, we can choose $F(u) = -e^{-u}$. Then

$$\begin{aligned} \int_0^{\ln 2} e^{-u} du &= F(\ln 2) - F(0) \\ &= -e^{-\ln 2} - (-e^0) = -(1/2) - (-1) = 1/2 \end{aligned}$$

Example. Evaluate $\int_2^3 \ln t dt$.

For the antiderivative, we can choose $F(t) = t \ln t - t$. Using the product rule, we can show that this is indeed an antiderivative:

$$F'(t) = t \cdot \frac{1}{t} + 1 \cdot \ln t - 1 = 1 + \ln t - 1 = \ln t$$

Thus,

$$\begin{aligned} \int_2^3 \ln t dt &= F(3) - F(2) \\ &= 3 \cdot \ln 3 - 3 - (2 \cdot \ln 2 - 2) \\ &= \ln 3^3 - \ln 2^2 - 1 = \ln 27 - \ln 4 - 1 = \ln(27/4) - 1 \end{aligned}$$

While formulas make it possible to get exact values, they do present us with problems of their own. For instance, we need to know that $t \ln t - t$ is an antiderivative of $\ln t$. This is not obvious. In fact, there is no guarantee that the antiderivative of a function given by a formula will have a formula! The antiderivatives of $\cos(x^2)$ and $\sin(x)/x$ do not have formulas, for instance. Many techniques have been devised to find the formula for an antiderivative. In chapter 11 of Calculus II we will survey some of those that are most frequently used.

Parameters

In chapter 4.2 we considered differential equations that involved parameters (see pages 214–218). It also happens that integrals can involve parameters. However, parameters complicate numerical work. If we calculate the value of an integral numerically, by making estimates with Riemann sums, we must first fix the value of any parameters that appear. This makes it difficult to

When integrals depend on parameters ...

see how the value of the integral depends on the parameters. We would have to give new values to the parameters and then recalculate the Riemann sums.

... the fundamental theorem can help make the relation clearer

The outcome is much simpler and more transparent if we are able to use the fundamental theorem to get a *formula* for the integral. The parameters just appear in the formula, so it is immediately clear how the integral depends on the parameters.

Here is an example that we shall explore further later. We want to see how the integrals

$$\int_a^b \sin(\alpha x) dx \quad \text{and} \quad \int_a^b \cos(\alpha x) dx$$

depend on the parameter $\alpha \neq 0$, and also on the parameters a and b . To begin, you should check that $F(x) = -\cos(\alpha x)/\alpha$ is an antiderivative of $\sin(\alpha x)$. Therefore,

$$\int_a^b \sin(\alpha x) dx = \frac{-\cos(\alpha b)}{\alpha} - \frac{-\cos(\alpha a)}{\alpha} = \frac{\cos(\alpha a) - \cos(\alpha b)}{\alpha}.$$

In a similar way, you should be able to show that

$$\int_a^b \cos(\alpha x) dx = \frac{\sin(\alpha b) - \sin(\alpha a)}{\alpha}.$$

Suppose the interval $[a, b]$ is exactly one-half of a full period: $[0, \pi/\alpha]$. Then

$$\begin{aligned} \int_0^{\pi/\alpha} \sin(\alpha x) dx &= \frac{\cos(\alpha \cdot 0) - \cos(\alpha \pi/\alpha)}{\alpha} \\ &= \frac{\cos 0 - \cos \pi}{\alpha} = \frac{1 - (-1)}{\alpha} = \frac{2}{\alpha}. \end{aligned}$$

Exercises

Constructing accumulation functions

- a) Obtain a formula for the accumulation function

$$A(X) = \int_2^X 5 dx$$

and sketch its graph on the interval $2 \leq X \leq 6$.

- b) Is $A'(X) = 5$?

2. Let $f(x) = 2 + x$ on the interval $0 \leq x \leq 5$.

a) Sketch the graph of $y = f(x)$.

b) Obtain a formula for the accumulation function

$$A(X) = \int_0^X f(x) dx$$

and sketch its graph on the interval $0 \leq X \leq 5$.

c) Verify that $A'(X) = f(X)$ for every X in $0 \leq X \leq 5$.

d) By comparing the graphs of f and A , verify that, at any point X , the *slope* of the graph of A is the same as the *height* of the graph of f .

3. a) Consider the accumulation function

$$A(X) = \int_0^X x^3 dx.$$

Using the fact that $A'(X) = X^3$, obtain a formula that expresses A in terms of X .

b) Modify A so that accumulation begins at the value $x = 1$ instead of $x = 0$ as in part (a). Thus

$$A(X) = \int_1^X x^3 dx.$$

It is still true that $A'(X) = X^3$, but now $A(1) = 0$. Obtain a formula that expresses this modified A in terms of X . How do the formulas for A in parts (a) and (b) differ?

Using the fundamental theorem

4. Find $A'(X)$ when

a) $A(X) = \int_0^X \cos(x) dx$

e) $A(X) = \int_0^X \sin(x^2) dx$

b) $A(X) = \int_0^X \sin(x) dx$

f) $A(X) = \int_0^X \sin^2 x dx$

c) $A(X) = \int_0^X \cos(x^2) dx$

g) $A(X) = \int_1^X \ln t dt$

d) $A(X) = \int_0^X \cos(t^2) dt$

h) $A(X) = \int_0^X x^2 - 4x^3 dx$

5. Find all critical points of the function

$$A(X) = \int_0^X \cos(x^2) dx$$

on the interval $0 \leq X \leq 4$. Indicate which critical points are local maxima and which are local minima. (Critical points and local maxima and minima are discussed on pages 303–309.)

[Answer: There are five critical points in the interval $[0, 4]$. The first is a local maximum at $\sqrt{\pi/2}$.]

6. Find all critical points of the function

$$A(X) = \int_0^X \sin(x^2) dx$$

on the interval $0 \leq X \leq 4$. Indicate which critical points are local maxima and which are local minima.

7. Find *all* critical points of the function

$$A(X) = \int_0^X x^2 - 4x^3 dx.$$

Indicate which critical points are local maxima and which are local minima.

8. Express the solution to each of the following initial value problems as an accumulation function (that is, as an integral with a variable upper limit of integration).

$$\begin{array}{ll} \text{a) } y' = \cos(x^2), & y(\sqrt{\pi}) = 0 \\ \text{b) } y' = \sin(x^2), & y(0) = 0 \end{array} \quad \begin{array}{ll} \text{c) } y' = \sin(x^2), & y(0) = 5 \\ \text{d) } y' = e^{-x^2}, & y(0) = 0 \end{array}$$

9. Sketch the graphs of the following accumulation functions over the indicated intervals.

$$\begin{array}{l} \text{a) } \int_0^X \sin(x^2) dx, \quad 0 \leq X \leq 4 \\ \text{b) } \int_0^X \frac{\sin(x)}{x} dx, \quad 0 \leq X \leq 4 \end{array}$$

Formulas for integrals

10. Determine the exact value of each of the following integrals.

a) $\int_3^7 2 - 3x + 5x^2 dx$

e) $\int_1^6 dx/x$

b) $\int_0^{5\pi} \sin x dx$

f) $\int_0^4 7u - 12u^5 du$

c) $\int_0^{5\pi} \sin(2x) dx$

g) $\int_0^1 2^t dt$

d) $\int_0^1 e^t dt$

h) $\int_{-1}^1 s^2 ds$

11. Express the values of the following integrals in terms of the parameters they contain.

a) $\int_3^7 kx dx$

e) $\int_{\ln 2}^{\ln 3} e^{ct} dt$

b) $\int_0^\pi \sin(\alpha x) dx$

f) $\int_1^b 5 - x dx$

c) $\int_1^4 px^2 - x^3 dx$

g) $\int_0^1 a^t dt$

d) $\int_0^1 e^{ct} dt$

h) $\int_1^2 u^c du$

12. Find a formula for the solution of each of the following initial value problems.

a) $y' = x^2 - 4x^3, \quad y(0) = 0$

c) $y' = x^2 - 4x^3, \quad y(3) = 10$

b) $y' = x^2 - 4x^3, \quad y(3) = 0$

d) $y' = \cos(3x), \quad y(\pi) = 0$

13. Find the average value of each of the following functions over the indicated interval.

a) $x^2 - x^3$ over $[0, 1]$

b) $\ln x$ over $[1, e]$

c) $\sin x$ over $[0, \pi]$

14. a) What is the average value of the function $px - x^2$ on the interval $[0, 1]$? The average depends on the parameter p .
- b) For which value of p will that average be zero?

6.5 Chapter Summary

The Main Ideas

- A **Riemann sum** for the function $f(x)$ on the interval $[a, b]$ is a sum of the form

$$f(x_1) \cdot \Delta x_1 + f(x_2) \cdot \Delta x_2 + \cdots + f(x_n) \cdot \Delta x_n,$$

where the interval $[a, b]$ has been subdivided into n subintervals whose lengths are $\Delta x_1, \Delta x_2, \dots, \Delta x_n$, and each x_k is a sampling point in the k -th subinterval (for each k from 1 to n).

- Riemann sums can be used to approximate a variety of quantities expressed as **products** where one factor varies with the other.
- Riemann sums give more accurate **approximations** as the lengths $\Delta x_1, \Delta x_2, \dots, \Delta x_n$ are made small.
- If the Riemann sums for a function $f(x)$ on an interval $[a, b]$ converge, the limit is called the **integral** of $f(x)$ on $[a, b]$, and it is denoted

$$\int_a^b f(x) dx.$$

- The **units** of $\int_a^b f(x) dx$ equal the product of the units of $f(x)$ and the units of x .
- **The Fundamental Theorem of Calculus.**
The solution $y = A(x)$ of the initial value problem

$$y' = f(x) \quad y(a) = 0$$

is the **accumulation function**

$$A(X) = \int_a^X f(x) dx.$$

- If $F(x)$ is an **antiderivative** of $f(x)$, then

$$\int_a^b f(x) dx = F(b) - F(a).$$

- The integral $\int_a^b f(x) dx$ equals the **signed area** between the graph of $f(x)$ and the x -axis.
- If $f(x)$ is **monotonic** on $[a, b]$ and if $\int_a^b f(x) dx$ is approximated by a Riemann sum with subintervals of width Δx , then the **error** in the approximation is at most $\Delta x \cdot |f(b) - f(a)|$.

Self-Testing

- You should be able to write down (by hand) a Riemann sum to approximate a quantity expressed as a product (e.g., human effort, electrical energy, work, distance travelled, area).
- You should be able to write down an integral giving the *exact* value of a quantity approximated by a Riemann sum.
- You should be able to use **sigma notation** to abbreviate a sum, and you should be able to read sigma notation to calculate a sum.
- You should be able to use a computer program to compute the value of a Riemann sum.
- You should be able to find an error bound when approximating an integral by a Riemann sum.
- You should know and be able to use the **integration rules**.
- You should be able to use the fundamental theorem of calculus to find the value of an integral.
- You should be able to use an antiderivative to find the value of an integral.

