4.2: Riemann Sums

Skills to Develop

In this section, we strive to understand the ideas generated by the following important questions:

- How can we use a Riemann sum to estimate the area between a given curve and the horizontal axis over a particular interval?
- What are the differences among left, right, middle, and random Riemann sums?
- How can we write Riemann sums in an abbreviated form??

In Section 4.1, we learned that if we have a moving object with velocity function \(v\), whenever \(v(t)\) is positive, the area between \(y = v(t)\) and the t-axis over a given time interval tells us the distance traveled by the object over that time period; in addition, if \(v(t)\) is sometimes negative and we view the area of any region below the t-axis as having an associated negative sign, then the sum of these signed areas over a given interval tells us the moving object’s change in position over the time interval. For instance, for the velocity function given in Figure \(\PageIndex{1}\), if the areas of shaded regions are \(A_1\), \(A_2\), and \(A_3\) as labeled, then the total distance \(D\) traveled by the moving object on \([a, b]\) is

\[D = A_1 + A_2 + A_3,\]

while the total change in the object’s position on \([a, b]\) is

\[s(b) - s(a) = A_1 - A_2 + A_3.\]

Because the motion is in the negative direction on the interval where \(v(t) < 0\), we subtract \(A_2\) when determining the object’s total change in position.
Figure (PageIndex{1}): A velocity function that is sometimes negative.

Of course, finding \( \Delta(D) \) and \( s(b) - s(a) \) for the situation given in Figure (PageIndex{1}) presumes that we can actually find the areas represented by A1, A2, and A3. In most of our work in Section 4.1, such as in Activities 4.2 and 4.3, we worked with velocity functions that were either constant or linear, so that by finding the areas of rectangles and triangles, we could find the area bounded by the velocity function and the horizontal axis exactly. But when the curve that bounds a region is not one for which we have a known formula for area, we are unable to find this area exactly. Indeed, this is one of our biggest goals in Chapter 4: to learn how to find the exact area bounded between a curve and the horizontal axis for as many different types of functions as possible. To begin, we expand on the ideas in Activity 4.1, where we encountered a nonlinear velocity function and approximated the area under the curve using four and eight rectangles, respectively. In the following preview activity, we focus on three different options for deciding how to find the heights of the rectangles we will use.

Preview Activity (PageIndex{1})

A person walking along a straight path has her velocity in miles per hour at time \( t \) given by the function \( v(t) = 0.25t^3 - 1.5t^2 + 3t + 0.25 \), for times in the interval \( 0 \leq t \leq 2 \). The graph of this function is also given in each of the three diagrams in Figure (PageIndex{2}). Note that in each diagram, we use four rectangles to estimate the area under \( y = v(t) \) on the interval \([0, 2] \), but the method by which the four rectangles’ respective heights are decided varies among the three individual graphs.

Figure (PageIndex{2}): Three approaches to estimating the area under \( y = v(t) \) on the interval \([0, 2] \).
How are the heights of rectangles in the left-most diagram being chosen? Explain, and hence determine the value of \[S = A_1 + A_2 + A_3 + A_4\] by evaluating the function \(y = v(t)\) at appropriately chosen values and observing the width of each rectangle. Note, for example, that \(A_3 = v(1) \cdot \frac{1}{2} = 2 \cdot \frac{1}{2} = 1\).

a. Explain how the heights of rectangles are being chosen in the middle diagram and find the value of \[T = B_1 + B_2 + B_3 + B_4\].

b. Likewise, determine the pattern of how heights of rectangles are chosen in the right-most diagram and determine \(U = C_1 + C_2 + C_3 + C_4\).

c. Of the estimates \(S\), \(T\), and \(U\), which do you think is the best approximation of \(D\), the total distance the person traveled on \([0, 2]\)? Why?

---

**Sigma Notation**

It is apparent from several different problems we have considered that sums of areas of rectangles is one of the main ways to approximate the area under a curve over a given interval. Intuitively, we expect that using a larger number of thinner rectangles will provide a way to improve the estimates we are computing. As such, we anticipate dealing with sums with a large number of terms. To do so, we introduce the use of so-called sigma notation, named for the Greek letter Σ, which is the capital letter S in the Greek alphabet. For example, say we are interested in the sum \(1 + 2 + 3 + \ldots + 100\), which is the sum of the first 100 natural numbers. Sigma notation provides a shorthand notation that recognizes the general pattern in the terms of the sum. It is equivalent to write

\[
\sum_{k=1}^{100} k = 1 + 2 + 3 + \ldots + 100.
\]

We read the symbol

\[
\sum_{k=1}^{100}
\]

as “the sum from \(k\) equals 1 to 100 of \(k\).” The variable \((k)\) is usually called the *index of summation*, and the letter that is used for this variable is immaterial. Each sum in sigma notation involves a function of the index; for example,

\[
\sum_{k=1}^{10} (k^2 + 2k) = (1^2 + 2 \cdot 1) + (2^2 + 2 \cdot 2) + (3^2 + 2 \cdot 3) + \ldots + (10^2 + 2 \cdot 10),
\]

and more generally,

\[
\sum_{k=1}^{n} f(k) = f(1) + f(2) + \ldots + f(n).
\]

Sigma notation allows us the flexibility to easily vary the function being used to track the pattern in the sum, as well as to adjust the number of terms in the sum simply by changing the value of \(n\). We test our understanding of this new notation in the following activity.

**Activity**

For each sum written in sigma notation, write the sum long-hand and evaluate the sum to find its value. For each sum written
in expanded form, write the sum in sigma notation.

a. \( \sum_{k=1}^{5} (k^2 + 2) \)

b. \( \sum_{i=3}^{6} (2i - 1) \)

c. \( 3 + 7 + 11 + 15 + \ldots + 27 \)

d. \( 4 + 8 + 16 + 32 + \ldots + 256 \)

e. \( \sum_{i=1}^{6} \frac{1}{2^i} \)

---

### Riemann Sums

When a moving body has a positive velocity function \( y = v(t) \) on a given interval \([a, b]\), we know that the area under the curve over the interval is the total distance the body travels on \([a, b]\). While this is the fundamental motivating force behind our interest in the area bounded by a function, we are also interested more generally in being able to find the exact area bounded by \( y = f(x) \) on an interval \([a, b]\), regardless of the meaning or context of the function \( f \). For now, we continue to focus on determining an accurate estimate of this area through the use of a sum of the areas of rectangles, doing so in the setting where \( f(x) \geq 0 \) on \([a, b]\). Throughout, unless otherwise indicated, we also assume that \( f \) is continuous on \([a, b]\). The first choice we make in any such approximation is the number of rectangles. If we say that the total number of rectangles is \( n \), and we desire \( n \) rectangles of equal width to subdivide the interval \([a, b]\), then each rectangle must have width \( 4x = \frac{b-a}{n} \). We observe further that

\[
x_1 = x_0 + 4x, \quad x_2 = x_0 + 24x,
\]

and thus in general

\[
x_i = a + i4x
\]

as pictured in Figure \( \PageIndex{3} \). We use each subinterval \([x_i, x_{i+1}]\) as the base of a rectangle, and next must choose how to decide the height of the rectangle that will be used to approximate the area under \( y = f(x) \) on the subinterval. There are three standard choices: use the left endpoint of each subinterval, the right endpoint of each subinterval, or the midpoint of each. These are precisely the options encountered in Preview Activity 4.2 and seen in Figure \( \PageIndex{2} \). We next explore how these choices can be reflected in sigma notation. If we now consider an arbitrary positive function \( f \) on \([a, b]\) with the interval subdivided as shown in Figure \( \PageIndex{3} \), and choose to use left endpoints, then on each interval of the form \([x_i, x_{i+1}]\), the area of the rectangle formed is given by \( A_{i+1} = f(x_i) \cdot 4x \), as seen in Figure \( \PageIndex{4} \). If we let \( L_n \) denote the sum of the areas of rectangles whose heights are given by the function value at each respective left endpoint, then we see that
\[ \sum_{i=0}^{n-1} f(x_i) \cdot 4x \]

In the more compact sigma notation, we have \( \sum_{i=0}^{n-1} f(x_i) \cdot 4x \). Note particularly that since the index of summation begins at 0 and ends at \( n - 1 \), there are indeed \( n \) terms in this sum. We call \( \sum_{i=0}^{n-1} f(x_i) \cdot 4x \) the left Riemann sum for the function \( f \) on the interval \([a, b]\).

**Figure 4:** Subdividing the interval \([a, b]\) into \( n \) subintervals of equal length \( 4x \) and approximating the area under \( y = f(x) \) over \([a, b]\) using left rectangles.

There are now two fundamental issues to explore: the number of rectangles we choose to use and the selection of the pattern by which we identify the height of each rectangle. It is best to explore these choices dynamically, and the applet found at http://gvsu.edu/s/a9 is a particularly useful one. There we see the image shown in

**Figure 5:** A snapshot of the applet found at http://gvsu.edu/s/a9.

Figure 5, but with the opportunity to adjust the slider bars for the left endpoint and the number of subintervals. By moving the sliders, we can see how the heights of the rectangles change as we consider left endpoints, midpoints, and right endpoints, as well as the impact that a larger number of narrower rectangles has on the approximation of the exact area bounded by the function and the horizontal axis. To see how the Riemann sums for right endpoints and midpoints are constructed, Marc Renault, Geogebra Calculus Applets. we consider Figure 6. For the sum with right endpoints, we see that the area of the
rectangle on an arbitrary interval \([x_i, x_{i+1}]\) is given by \(B_{i+1} = f(x_{i+1}) \cdot 4x\), so that the sum of all such areas of rectangles is given by

\[
R_n = B_1 + B_2 + \ldots + B_{i+1} + \ldots + B_n = f(x_1) \cdot 4x + f(x_2) \cdot 4x + \ldots + f(x_{i+1}) \cdot 4x + \ldots + f(x_n) \cdot 4x = \sum_{i=1}^{n} f(x_i) \cdot 4x.
\]

We call \(R_n\) the right Riemann sum for the function \(f\) on the interval \([a, b]\). For the sum that uses midpoints, we introduce the notation \(x_{i+1} = x_i + x_{i+1}/2\) so that \(x_{i+1}\) is the midpoint of the interval \([x_i, x_{i+1}]\). For instance, for the rectangle with area \(C_1\) in Figure \(\PageIndex{6}\), we now have \(C_1 = f(x_1) \cdot 4x\). Hence, the sum of all the areas of rectangles that use midpoints is

\[
M_n = C_1 + C_2 + \ldots + C_{i+1} + \ldots + C_n = f(x_1) \cdot 4x + f(x_2) \cdot 4x + \ldots + f(x_{i+1}) \cdot 4x + \ldots + f(x_n) \cdot 4x = \sum_{i=1}^{n} f(x_i) \cdot 4x,
\]

and we say that \(M_n\) is the middle Riemann sum for \(f\) on \([a, b]\). When \(f(x) \geq 0\) on \([a, b]\), each of the Riemann sums \(L_n, R_n,\) and \(M_n\) provides an estimate of the area under the curve \(y = f(x)\) over the interval \([a, b]\); momentarily, we will discuss the meaning of Riemann sums in the setting when \(f\) is sometimes negative. We also recall that in the context of a nonnegative velocity function \(v = v(t)\), the corresponding Riemann sums are approximating the distance traveled on \([a, b]\) by the moving object with velocity function \(v\). There is a more general way to think of Riemann sums, and that is to not restrict the choice of where the function is evaluated to determine the respective rectangle heights. That is, rather than saying we’ll always choose left endpoints, or always choose midpoints, we simply say that a point \(x_i \in (x_{i-1}, x_i)\) (so that \(x_i \leq x \leq x_i + 1\)), which makes the Riemann sum given by \(f(x_i) \cdot 4x + f(x_{i+1}) \cdot 4x + \ldots + f(x_n) \cdot 4x = \sum_{i=1}^{n} f(x_i) \cdot 4x\). At http://gvsu.edu/s/a9, the applet noted earlier and referenced in Figure \(\PageIndex{5}\), by unchecking the “relative” box at the top left, and instead checking “random,” we can easily explore the effect of using random point locations in subintervals on a given Riemann sum. In computational practice, we most often use \(L_n, R_n,\) or \(M_n\), while the random Riemann sum is useful in theoretical discussions. In the following activity, we investigate several different Riemann sums for a particular velocity function.

Activity \(\PageIndex{3}\)

Suppose that an object moving along a straight line path has its velocity in feet per second at time \(t\) in seconds given by \(v(t) = 29(t - 3)^2 + 2\).

a. Carefully sketch the region whose exact area will tell you the value of the distance the object traveled on the time interval \(2 \leq t \leq 5\).

b. Estimate the distance traveled on \([2, 5]\) by computing \(L_4, R_4,\) and \(M_4\).
c. Does averaging L4 and R4 result in the same value as M4? If not, what do you think the average of L4 and R4 measures?

d. For this question, think about an arbitrary function $f$, rather than the particular function $v$ given above. If $f$ is positive and increasing on $[a, b]$, will $L_n$ overestimate or under-estimate the exact area under $f$ on $[a, b]$? Will $R_n$ over- or under-estimate the exact area under $f$ on $[a, b]$? Explain.

When the function is sometimes negative

For a Riemann sum such as $L_n = \sum_{i=0}^{n-1} f(x_i)\Delta x$, we can of course compute the sum even when $f$ takes on negative values. We know that when $f$ is positive on $[a, b]$, the corresponding left Riemann sum $L_n$ estimates the area bounded by $f$ and the horizontal axis over the interval. For a function such as the

![Graph showing Riemann sums for a function that is sometimes negative.](image)

Figure \(\PageIndex{7}\): At left and center, two left Riemann sums for a function $f$ that is sometimes negative; at right, the areas bounded by $f$ on the interval $[a, d]$.

one pictured in Figure \(\PageIndex{7}\), where in the first figure a left Riemann sum is being taken with 12 subintervals over $[a, d]$, we observe that the function is negative on the interval $b \leq x \leq c$, and so for the four left endpoints that fall in $[b, c]$, the terms $f(x_i)\Delta x$ have negative function values. This means that those four terms in the Riemann sum produce an estimate of the opposite of the area bounded by $y = f(x)$ and the x-axis on $[b, c]$. In Figure \(\PageIndex{7}\), we also see evidence that by increasing the number of rectangles used in a Riemann sum, it appears that the approximation of the area (or the opposite of the area) bounded by a curve appears to improve. For instance, in the middle graph, we use 24 left rectangles, and from the shaded areas, it appears that we have decreased the error from the approximation that uses 12. When we proceed to Section 4.3, we will discuss the natural idea of letting the number of rectangles in the sum increase without bound. For now, it is most important for us to observe that, in general, any Riemann sum of a continuous function $f$ on an interval $[a, b]$ approximates the difference between the area that lies above the horizontal axis on $[a, b]$ and under $f$ and the area that lies below the horizontal axis on $[a, b]$ and above $f$. In the notation of Figure \(\PageIndex{7}\), we may say that

$$L_{24} \approx A_1 - A_2 + A_3,$$

where $L_{24}$ is the left Riemann sum using 24 subintervals shown in the middle graph, and $A_1$ and $A_3$ are the areas of the regions where $f$ is positive on the interval of interest, while $A_2$ is the area of the region where $f$ is negative. We will also call the quantity $A_1 - A_2 + A_3$ the net signed area bounded by $f$ over the interval $[a, d]$, where by the phrase “signed area” we indicate that we are attaching a minus sign to the areas of regions that fall below the horizontal axis.

Finally, we recall from the introduction to this present section that in the context where the function $f$ represents the velocity of a moving object, the total sum of the areas bounded by the curve tells us the total distance traveled over the relevant time interval, while the total net signed area bounded by the curve computes the object’s change in position on the interval.

Activity \(\PageIndex{4}\)
Suppose that an object moving along a straight line path has its velocity \( v \) (in feet per second) at time \( t \) (in seconds) given by

\[
v(t) = \frac{1}{2} t^2 - 3t + \frac{7}{2}.
\]

a. Compute \( M_5 \), the middle Riemann sum, for \( v \) on the time interval \([1, 5]\). Be sure to clearly identify the value of \( 4t \) as well as the locations of \( t_0, t_1, \ldots, t_5 \). In addition, provide a careful sketch of the function and the corresponding rectangles that are being used in the sum.

b. Building on your work in (a), estimate the total change in position of the object on the interval \([1, 5]\).

c. Building on your work in (a) and (b), estimate the total distance traveled by the object on \([1, 5]\).

d. Use appropriate computing technology to compute \( M_{10} \) and \( M_{20} \). What exact value do you think the middle sum eventually approaches as \( n \) increases without bound? What does that number represent in the physical context of the overall problem?

**Summary**

In this section, we encountered the following important ideas:

- A Riemann sum is simply a sum of products of the form \( f(x^*_i) \Delta x \) that estimates the area between a positive function and the horizontal axis over a given interval. If the function is sometimes negative on the interval, the Riemann sum estimates the difference between the areas that lie above the horizontal axis and those that lie below the axis.

- The three most common types of Riemann sums are left, right, and middle sums, plus we can also work with a more general, random Riemann sum. The only difference is that the function is evaluated at different points within each subinterval, while for right and middle sums, we use right endpoints and midpoints, respectively.

- The left, right, and middle Riemann sums are denoted \( L_n \), \( R_n \), and \( M_n \), with formulas:

\[
L_n = f(x_0)\Delta x + f(x_1)\Delta x + \ldots + f(x_{n-1})\Delta x = \sum_{i=0}^{n-1} f(x_i)\Delta x,
\]

\[
R_n = f(x_1)\Delta x + f(x_2)\Delta x + \ldots + f(x_n)\Delta x = \sum_{i=1}^{n} f(x_i)\Delta x,
\]

\[
M_n = f(x_1)\Delta x + f(x_2)\Delta x + \ldots + f(x_n)\Delta x = \sum_{i=1}^{n} f(x_i)\Delta x, \quad x_i = (x_{i-1} + x_i)/2.
\]

**Contributors and Attributions**

Matt Boelkins (Grand Valley State University), David Austin (Grand Valley State University), Steve Schlicker (Grand Valley State University)