5.3: The Definite Integral

Skills to Develop

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

- How does increasing the number of subintervals affect the accuracy of the approximation generated by a Riemann sum?
- What is the definition of the definite integral of a function $f$ over the interval $[a, b]$?
- What does the definite integral measure exactly, and what are some of the key properties of the definite integral?

In Figure 1, we see visual evidence that increasing the number of rectangles in a Riemann sum improves the accuracy of the approximation of the net signed area that is bounded by the given function on the interval under consideration. We thus explore the natural idea of allowing the number of rectangles to increase without bound in an effort to compute the exact net signed area bounded by a function on an interval. In addition, it is important to think about the differences among left, right, and middle Riemann sums and the different results they generate as the value of $n$ increases. As we have done throughout our investigations with area, we begin with functions that are exclusively positive on the interval under consideration.

*Figure 1*: At left and center, two left Riemann sums for a function $f$ that is sometimes negative; at right, the exact areas bounded by $f$ on the interval $[a, d]$. 
Preview Activity

Consider the applet found at http://gvsu.edu/s/aw6. There, you will initially see the situation shown in Figure 2. Note that the value of the chosen Riemann sum is displayed next to the word “relative,” and that you can change the type of Riemann sum being computed by dragging the point on the slider bar below the phrase “sample point placement.” Explore to see how you can change the window in which the function is viewed, as well as the function itself. You can set the minimum and maximum values of \(x\) by clicking and dragging on the blue points that set the endpoints; you can change the function by typing a new formula in the “\(f(x)\)” window at the bottom; and you can adjust the overall window by “panning and zooming” by using the Shift key and the scrolling feature of your mouse. More information on how to pan and zoom can be found at http://gvsu.edu/s/FI. Work accordingly to adjust the applet so that it uses a left Riemann sum with \(n = 5\) subintervals for the function is \(f(x) = 2x + 1\). You should see the updated figure shown in Figure 3.

\[ f(x) = \sin(2x) - \frac{x^2}{10} + 3 \]

**Figure 2:** A right Riemann sum with 10 subintervals for the function \(f(x) = \sin(2x) - x^2/10 + 3\) on the interval \([1, 7]\). The value of the sum is \(R_{10} = 4.90595\).

Riemann sum is displayed next to the word “relative,” and that you can change the type of Riemann sum being computed by dragging the point on the slider bar below the phrase “sample point placement.” Explore to see how you can change the window in which the function is viewed, as well as the function itself. You can set the minimum and maximum values of \(x\) by clicking and dragging on the blue points that set the endpoints; you can change the function by typing a new formula in the “\(f(x)\)” window at the bottom; and you can adjust the overall window by “panning and zooming” by using the Shift key and the scrolling feature of your mouse. More information on how to pan and zoom can be found at http://gvsu.edu/s/FI. Work accordingly to adjust the applet so that it uses a left Riemann sum with \(n = 5\) subintervals for the function is \(f(x) = 2x + 1\). You should see the updated figure shown in Figure 3.

\[ f(x) = 2x + 1 \]

**Figure 3:** A left Riemann sum with 5 subintervals for the function \(f(x) = 2x + 1\) on the interval \([1, 4]\). The value of the sum is \(L_5 = 16.2\).

Then, answer the following questions.
a. Update the applet (and view window, as needed) so that the function being considered is \( f(x) = 2x + 1 \) on \([1, 4]\), as directed above. For this function on this interval, compute \( L_n, M_n, R_n \) for \( n = 5, n = 25, \) and \( n = 100 \). What appears to be the exact area bounded by \( f(x) = 2x + 1 \) and the x-axis on \([1, 4]\)?

b. Use basic geometry to determine the exact area bounded by \( f(x) = 2x + 1 \) and the x-axis on \([1, 4]\).

c. Based on your work in (a) and (b), what do you observe occurs when we increase the number of subintervals used in the Riemann sum?

d. Update the applet to consider the function \( f(x) = x^2 + 1 \) on the interval \([1, 4]\) (note that you need to enter “\( x^2 + 1 \)” for the function formula). Use the applet to compute \( L_n, M_n, R_n \) for \( n = 5, n = 25, \) and \( n = 100 \). What do you conjecture is the exact area bounded by \( f(x) = x^2 + 1 \) and the x-axis on \([1, 4]\)?

e. Why can we not compute the exact value of the area bounded by \( f(x) = x^2 + 1 \) and the x-axis on \([1, 4]\) using a formula like we did in (b)?

The Definition of the Definite Integral

In both examples in Preview Activity \( \PageIndex{1} \), we saw that as the number of rectangles got larger and larger, the values of \( L_n, M_n, \) and \( R_n \) all grew closer and closer to the same value. It turns out that this occurs for any continuous function on an interval \([a, b]\), and even more generally for a Riemann sum using any point \( x^?_{i+1} \) in the interval \([x_i, x_{i+1}]\). Said differently, as we let \( n \to \infty \), it doesn’t really matter where we choose to evaluate the function within a given subinterval, because

\[
\lim_{n\to\infty} L_n = \lim_{n\to\infty} R_n = \lim_{n\to\infty} M_n = \lim_{n\to\infty} \sum^{n}_{i=1} f(x^?_i) \Delta x.
\]

That these limits always exist (and share the same value) for a continuous function \( f \) allows us to make the following definition.

Definition: Definite Integral

The definite integral of a continuous function \( f \) on the interval \([a, b]\), denoted \( \int_a^b f(x) \, dx \), is the real number given by

\[
\int_a^b f(x) \, dx = \lim_{n\to\infty} \sum^{n}_{i=1} f(x^?_i) \Delta x,
\]

where \( \Delta x = \frac{b-a}{n} \), \( x_i = a + i \Delta x \) (for \( i = 0, \ldots, n \)), and \( x^?_i \) satisfies \( x_{i-1} \leq x^?_i \leq x_i \) (for \( i = 1, \ldots, n \)).

We call the symbol \( \int \) the integral sign, the values \( a \) and \( b \) the limits of integration, and the function \( f \) the integrand. The process of determining the real number \( \int_a^b f(x) \, dx \) is called evaluating the definite integral. While we will come to understand that there are several different interpretations of the value of the definite integral, for now the most important is that

\[
\int_a^b f(x) \, dx \]

measures the net signed area bounded by \( y = f(x) \) and the x-axis on the interval \([a, b]\). For example, in the notation of the definite integral, if \( f \) is the function pictured in Figure \( \PageIndex{4} \) and \( A_1, A_2, \) and \( A_3 \) are the exact areas bounded by \( f \) and the x-axis on the respective intervals \([a, b], [b, c], \) and \([c, d]\), then

\[
\int_a^b f(x) \, dx = A_1
\]
\[\int^b_c f(x) \, dx = -A_2\]
\[\int^d_c f(x) \, dx = A_3\]
and
\[\int^d_a f(x) \, dx = A_1 - A_2 + A_3\]

We can also use definite integrals to express the change in position and distance traveled by a moving object. In the setting of a velocity function \(v\) on an interval \([a, b]\), it follows from our work above and in preceding sections that the change in position, \((s(b) - s(a))\), is given by

\[s(b) - s(a) = \int_a^b v(t) \, dt\]

It turns out that a function need not be continuous in order to have a definite integral. For our purposes, we assume that the functions we consider are continuous on the interval(s) of interest. It is straightforward to see that any function that is piecewise continuous on an interval of interest will also have a well-defined definite integral.

If the velocity function is nonnegative on \([a, b]\), then \(\int_a^b v(t) \, dt\) tells us the distance the object traveled. When velocity is sometimes negative on \([a, b]\), the areas bounded by the function on intervals where \(v\) does not change sign can be found using integrals, and the sum of these values will tell us the distance the object traveled. If we wish to compute the value of a definite integral using the definition, we have to take the limit of a sum. While this is possible to do in select circumstances, it is also tedious and time-consuming; moreover, computing these limits does not offer much additional insight into the meaning or interpretation of the definite integral. Instead, in Section 4.4, we will learn the Fundamental Theorem of Calculus, a result that provides a shortcut for evaluating a large class of definite integrals. This will enable us to determine the exact net signed area bounded by a continuous function and the x-axis in many circumstances, including examples such as \(\int 4 \, dx\), which we approximated by Riemann sums in Preview Activity. For now, our goal is to understand the
meaning and properties of the definite integral, rather than how to actually compute its value using ideas in calculus. Thus, we temporarily rely on the net signed area interpretation of the definite integral and observe that if a given curve produces regions whose areas we can compute exactly through known area formulas, we can thus compute the exact value of the integral. For instance, if we wish to evaluate the definite integral \( \int_4^1 (2x + 1) \, dx \), we can observe that the region bounded by this function and the x-axis is the trapezoid shown in Figure \( \PageIndex{5} \), and by the known formula for the area of a trapezoid, its area is \( A = \frac{1}{2} (3 + 9) \cdot 3 = 18 \), so \( \int_4^1 (2x + 1) \, dx = 18 \).

Activity \( \PageIndex{2} \)

Use known geometric formulas and the net signed area interpretation of the definite integral to evaluate each of the definite integrals below.

a. \( \int_1^0 3x \, dx \)

b. \( \int_{-1}^4 (2 - 2x) \, dx \)

c. \( \int_{-1}^1 \sqrt{1 - x^2} \, dx \)

d. \( \int_{-3}^4 \) where \( \{g\} \) is the function pictured in Figure \( \PageIndex{6} \). Assume that each portion of \( \{g\} \) is either part of a line or part of a circle.

Figure \( \PageIndex{5} \): The area bounded by \( f(x) = 2x + 1 \) and the x-axis on the interval \([1, 4]\).

Activity \( \PageIndex{2} \)

Use known geometric formulas and the net signed area interpretation of the definite integral to evaluate each of the definite integrals below.

a. \( \int_1^0 3x \, dx \)

b. \( \int_{-1}^4 (2 - 2x) \, dx \)

c. \( \int_{-1}^1 \sqrt{1 - x^2} \, dx \)

d. \( \int_{-3}^4 \) where \( \{g\} \) is the function pictured in Figure \( \PageIndex{6} \). Assume that each portion of \( \{g\} \) is either part of a line or part of a circle.

Figure \( \PageIndex{6} \): A function \( g \) that is piecewise defined; each piece of the function is part of a circle or part of a line.
Some Properties of the Definite Integral

With the perspective that the definite integral of a function \( f \) over an interval \([a, b]\) measures the net signed area bounded by \( f \) and the x-axis over the interval, we naturally arrive at several different standard properties of the definite integral. In addition, it is helpful to remember that the definite integral is defined in terms of Riemann sums that fundamentally consist of the areas of rectangles.

If we consider the definite integral \( R \ a \ a \ f (x) \, dx \) for any real number \( (a) \), it is evident that no area is being bounded because the interval begins and ends with the same point. Hence,

If \( f \) is a continuous function and \( a \) is a real number, then

\[
\int^a_a f (x) \, dx = 0. \]

Next, we consider the results of subdividing a given interval. In Figure \( \PageIndex{7} \), we see that

\[
\int_a^b f (x) \, dx = A_1
\]

\[
\int^c_b f (x) \, dx = A_2
\]

and

\[
\int^c_a f (x) \, dx = A_1 + A_2,
\]

which is indicative of the following general rule.

**Rule**

If \( f \) is a continuous function and \( a, b, \) and \( c \) are real numbers, then

\[
\int_Z^{c} f (x) \, dx = \int_a^{b} f (x) \, dx + \int^c_b f (x) \, dx. \]
While this rule is most apparent in the situation where \( a < b < c \), it in fact holds in general for any values of \( a, b, \) and \( c \). This result is connected to another property of the definite integral, which states that if we reverse the order of the limits of integration, we change the sign of the integral’s value. If \( \int f(x) \, dx \) is a continuous function and \( a \) and \( b \) are real numbers, then

\[
\int_a^b f(x) \, dx = - \int_b^a f(x) \, dx.
\]

This result makes sense because if we integrate from \( a \) to \( b \), then in the defining Riemann sum \( 4x = b - a \), while if we integrate from \( b \) to \( a \), \( 4x = a - b = - (b - a) \), and this is the only change in the sum used to define the integral. There are two additional properties of the definite integral that we need to understand. Recall that when we worked with derivative rules in Chapter 2, we found that both the Constant Multiple Rule and the Sum Rule held. The Constant Multiple Rule tells us that if \( f \) is a differentiable function and \( k \) is a constant, then

\[
\frac{d}{dx} [k f(x)] = k f'(x),
\]

and the Sum Rule states that if \( f \) and \( g \) are differentiable functions, then

\[
\frac{d}{dx} [f(x) + g(x)] = f'(x) + g'(x).
\]

These rules are useful because they enable us to deal individually with the simplest parts of certain functions and take advantage of the elementary operations of addition and multiplying by a constant. They also tell us that the process of taking the derivative respects addition and multiplying by constants in the simplest possible way.

It turns out that similar rules hold for the definite integral. First, let’s consider the situation pictured in Figure (PageIndex{8}), where we examine the effect of multiplying a function by a factor of 2 on the area it bounds with the x-axis. Because multiplying the function by 2 doubles its height at every x-value, we see that if we consider a typical rectangle from a Riemann sum, the difference in area comes from the changed height of the rectangle: \( f(x_i) \) for the original function, versus \( 2f(x_i) \) in the doubled function, in the case of left sum. Hence, in Figure (PageIndex{8}), we see that for the pictured rectangles with areas \( A \) and \( B \), it follows \( B = 2A \). As this will happen in every such rectangle, regardless of the value of \( n \) and the type of sum we use, we see that in the limit, the area of the red region bounded by \( y = 2f(x) \) will be twice that of the area of the blue region bounded by \( y = f(x) \). As there is nothing special about the value 2 compared to an arbitrary constant \( k \), it turns out that the following general principle holds.

Constant Multiple Rule

\[
\int_a^b k f(x) \, dx = k \int_a^b f(x) \, dx.
\]
If \( f \) is a continuous function and \( k \) is any real number then

\[
\int_a^b k \cdot f(x) \, dx = k \int_a^b f(x) \, dx.
\]

Finally, we see a similar situation geometrically with the sum of two functions \( f \) and \( g \). In particular, as shown in Figure \( \PageIndex{9} \), if we take the sum of two functions \( f \) and \( g \), at every point in the interval, the height of the function \( f+g \) is given by \( (f+g)(x) = f(x) + g(x) \), which is the sum of the individual function values of \( f \) and \( g \) (taken at left endpoints). Hence, for the pictured rectangles with areas \( A, B, \) and \( C \), it follows that \( C = A + B \), and because this will occur for every such rectangle, in the limit the area of the gray region will be the sum of the areas of the blue and red regions.

**Figure \( \PageIndex{9} \):** The areas bounded by \( y = f(x) \) and \( y = g(x) \) on \([a, b]\), as well as the area bounded by \( y = f(x) + g(x) \).

Stated in terms of definite integrals, we have the following general rule.

**Sum Rule**

If \( f \) and \( g \) are continuous functions, then

\[
\int_a^b [f(x) + g(x)] \, dx = \int_a^b f(x) \, dx + \int_a^b g(x) \, dx.
\]

More generally, the Constant Multiple and Sum Rules can be combined to make the observation that for any continuous functions \( f \) and \( g \) and any constants \( c \) and \( k \),

\[
\int_a^b [c f(x) \pm k g(x)] \, dx = c \int_a^b f(x) \, dx \pm k \int_a^b g(x) \, dx.
\]

**Activity \( \PageIndex{3} \)**

Suppose that the following information is known about the functions \( f \), \( g \), \( x^2 \), and \( x^3 \):

- \( \int_0^2 f(x) \, dx = -3; \int_2^5 f(x) \, dx = 2 \)
- \( \int_0^2 g(x) \, dx = 4; \int_2^5 g(x) \, dx = -1 \)
- \( \int_0^2 x^2 \, dx = 8.3; \int_2^5 x^2 \, dx = 117.3 \)
- \( \int_0^2 x^3 \, dx = 4; \int_2^5 x^3 \, dx = 609.4 \)

Use the provided information and the rules discussed in the preceding section to evaluate each of the following definite integrals.

a. \( \int_0^5 2 f(x) \, dx \)
b. \( \int_{0}^{5} g(x) \, dx \)

c. \( \int_{0}^{5} (f(x) + g(x)) \, dx \)

d. \( \int_{0}^{2} (3x^2 - 4x^3) \, dx \)

e. \( \int_{0}^{5} (2x^3 - 7g(x)) \, dx \)

---

**How the definite integral is connected to a function’s average value**

One of the most valuable applications of the definite integral is that it provides a way to meaningfully discuss the average value of a function, even for a function that takes on infinitely many values. Recall that if we wish to take the average of \( n \) numbers \( y_1, y_2, \ldots, y_n \), we do so by computing

\[
\text{Avg} = \frac{y_1 + y_2 + \cdots + y_n}{n}.
\]

Since integrals arise from Riemann sums in which we add \( n \) values of a function, it should not be surprising that evaluating an integral is something like averaging the output values of a function. Consider, for instance, the right Riemann sum \( R_n \) of a function \( f \), which is given by

\[
R_n = f(x_1)4x + f(x_2)4x + \cdots + f(x_n)4x = (f(x_1) + f(x_2) + \cdots + f(x_n))4x.
\]

Since \( 4x = b-a \), we can thus write

\[
R_n = (f(x_1) + f(x_2) + \cdots + f(x_n)) \cdot (b-a) = (b-a)f(x_1) + f(x_2) + \cdots + f(x_n) \cdot (4.1)
\]

Here, we see that the right Riemann sum with \( n \) subintervals is the length of the interval \( (b-a) \) times the average of the \( n \) function values found at the right endpoints. And just as with our efforts to compute area, we see that the larger the value of \( n \) we use, the more accurate our average of the values of \( f \) will be. Indeed, we will define the average value of \( f \) on \([a, b]\) to be

\[
f_{\text{AVG}[a,b]} = \lim_{n \to \infty} \frac{1}{b-a} \int_{a}^{b} f(x) \, dx.
\]

Solving Equation \ref{4.2} for \( f_{\text{AVG}[a,b]} \), we have the following general principle.

**Average value of a function:**

If \( f \) is a continuous function on \([a, b]\), then its average value on \([a, b]\) is given by the formula

\[
\left[ f_{\text{AVG}[a,b]} \right] = \frac{1}{b-a} \int_{a}^{b} f(x) \, dx.
\]

Observe that Equation \ref{4.2} tells us another way to interpret the definite integral: the definite integral of a function \( f \) from \( a \) to \( b \) is the length of the interval \( (b-a) \) times the average value of the function on the interval. In addition, Equation \ref{4.2} has a natural visual interpretation when the function \( f \) is nonnegative on \([a, b]\). Consider Figure \ref{PageIndex}, where we
see at left the shaded region whose area is $\int_a^b f(x) \, dx$, at center the shaded rectangle whose dimensions are $(b - a)$ by $f_{\text{AVG}}[a,b]$, and at right these two figures superimposed. Specifically, note that in dark green we show the horizontal line $y = f_{\text{AVG}}[a,b]$. Thus, the area of the green rectangle is given by $(b - a) \cdot f_{\text{AVG}}[a,b]$, which is precisely the value of $\int_a^b f(x) \, dx$. Said differently, the area of the blue region in the

![Figure](/figures.png): A function $y = f(x)$, the area it bounds, and its average value on $[a, b]$.

left figure is the same as that of the green rectangle in the center figure; this can also be seen by observing that the areas $A_1$ and $A_2$ in the rightmost figure appear to be equal. Ultimately, the average value of a function enables us to construct a rectangle whose area is the same as the value of the definite integral of the function on the interval. The java applet at [http://gvsu.edu/s/az](http://gvsu.edu/s/az) provides an opportunity to explore how the average value of the function changes as the interval changes, through an image similar to that found in Figure [PageIndex{10}].

Activity [PageIndex{4}]:

Suppose that $v(t) = \sqrt{4 - (t - 2)^2}$ tells us the instantaneous velocity of a moving object on the interval $0 \leq t \leq 4$, where $t$ is measured in minutes and $v$ is measured in meters per minute.

a. Sketch an accurate graph of $y = v(t)$. What kind of curve is $y = p 4 - (t - 2) 2$?
b. Evaluate $\int_0^4 v(t) \, dt$ exactly.
c. In terms of the physical problem of the moving object with velocity $v(t)$, what is the meaning of $\int_0^4 v(t) \, dt$? Include units on your answer.
d. Determine the exact average value of $v(t)$ on $[0, 4]$. Include units on your answer.
e. Sketch a rectangle whose base is the line segment from $t = 0$ to $t = 4$ on the t-axis such that the rectangle’s area is equal to the value of $\int_0^4 v(t) \, dt$. What is the rectangle’s exact height?
f. How can you use the average value you found in (d) to compute the total distance traveled by the moving object over $[0, 4]$? David Austin, [http://gvsu.edu/s/5r](http://gvsu.edu/s/5r).

**Summary**

In this section, we encountered the following important ideas:

- Any Riemann sum of a continuous function $f$ on an interval $[a, b]$ provides an estimate of the net signed area bounded by the function and the horizontal axis on the interval. Increasing the number of subintervals in the Riemann sum improves the accuracy of this estimate, and letting the number of subintervals increase without bound results in the values of the corresponding Riemann sums approaching the exact value of the enclosed net signed area.
- When we take the just described limit of Riemann sums, we arrive at what we call the definite integral of $f$ over the
interval \([a, b]\). In particular, the symbol \(\int_a^b f(x) \, dx\) denotes the definite integral of \(f\) over \([a, b]\), and this quantity is defined by the equation 
\[\int_a^b f(x) \, dx = \lim_{n \to \infty} \sum_{i=1}^{n} f(x_i^*) 4x,\]
where \(4x = \frac{b-a}{n}\), \(x_i = a + i4x\) (for \(i = 0, \ldots, n\)), and \(x_i^*\) satisfies \(x_{i-1} \leq x_i^* \leq x_i\) (for \(i = 1, \ldots, n\)).

- The definite integral \(\int_a^b f(x) \, dx\) measures the exact net signed area bounded by \(f\) and the horizontal axis on \([a, b]\); in addition, the value of the definite integral is related to what we call the average value of the function on \([a, b]\): 
  \(f_{\text{AVG}}[a,b] = \frac{1}{b-a} \cdot \int_a^b f(x) \, dx\). In the setting where we consider the integral of a velocity function \(v\), \(\int_a^b v(t) \, dt\) measures the exact change in position of the moving object on \([a, b]\); when \(v\) is nonnegative, \(\int_a^b v(t) \, dt\) is the object’s distance traveled on \([a, b]\).

- The definite integral is a sophisticated sum, and thus has some of the same natural properties that finite sums have. Perhaps most important of these is how the definite integral respects sums and constant multiples of functions, which can be summarized by the rule 
  \[\int_a^b [c f(x) \pm kg(x)] \, dx = c \int_a^b f(x) \, dx \pm k \int_a^b g(x) \, dx\]
  where \(f\) and \(g\) are continuous functions on \([a, b]\) and \(c\) and \(k\) are arbitrary constants.

---

**Contributors**

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