4.4 The Definite integral

Let f be a function that's non-negative on an interval $[a,b]$. Then, as we've seen in earlier sections, in a variety of circumstances, the area under the graph of f, over $[a,b]$, can be approximated by adding up areas of rectangles, whose bases lie on consecutive subintervals of $[a,b]$, and whose heights are given by values of f at "sampling points" x_k.

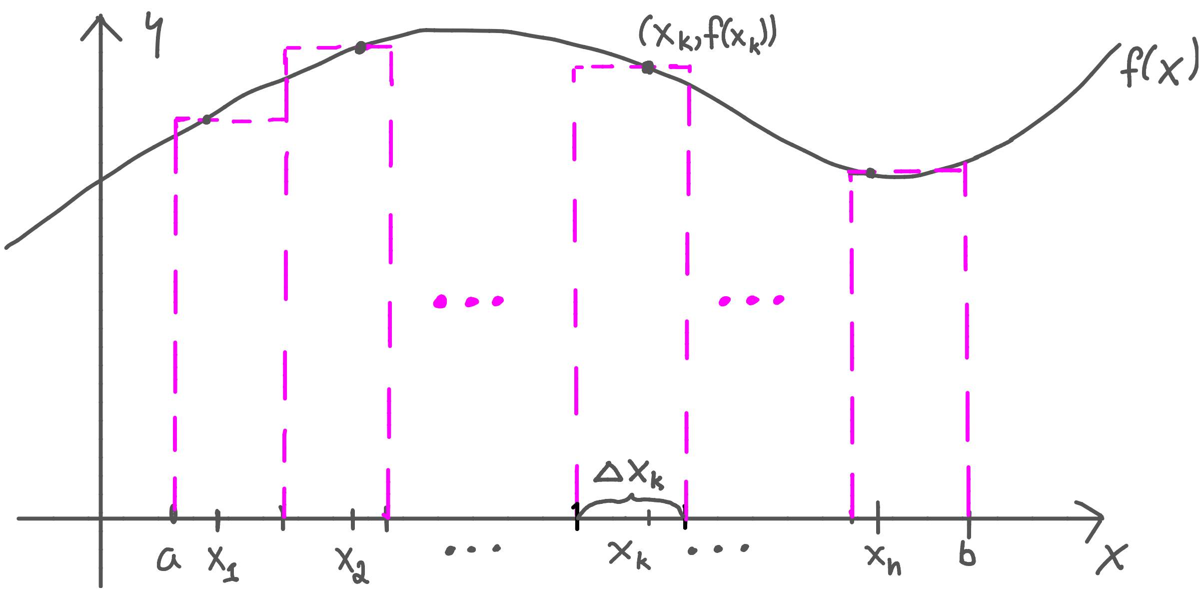


Figure 4.11. Approximating the area under a graph

That is, the area A under f, over this interval, can be approximated by a *Riemann sum:*

$$A \approx f(x_1)\Delta x_1 + f(x_2)\Delta x_2 + \cdots + f(x_n)\Delta x_n. \tag{4.4.1}$$

Intuitively, the more rectangles we use, and the narrower their baselengths, the closer these rectangles come to filling up the space under the graph precisely. So, at least intuitively, as the number n of rectangles increases to infinity, and all of the baselengths Δx_k shrink to zero, our Riemann sum should converge *exactly* to the area A in question.

When this happens, we give a special name, and a special symbol, to this area.

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

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

Equation (4.4.1) then tell us that

$$\int_a^b f(x)\,dx \approx f(x_1)\Delta x_1 + f(x_2)\Delta x_2 + \cdots + f(x_n)\Delta x_n \tag{4.4.2}$$

where, again, the right-hand side of (4.4.2) is a Riemann sum for f on $[a,b]$. But it tells us more: it says that, in fact,

$$\int_a^b f(x)\,dx = \lim_{\substack{\Delta x_k \to 0 \\ \text{for all } k}} \big(f(x_1)\Delta x_1 + f(x_2)\Delta x_2 + \cdots + f(x_n)\Delta x_n\big), \tag{4.4.3}$$

provided the conditions stated in the definition are met.

The function f is called the **integrand**. There *are* functions f whose Riemann sums *don't converge*, so that the integral in question *does not* exist. But such functions are rare. All the functions that typically arise in context, and nearly all the functions we study in calculus, *do* have integrals.

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 Δx_1, Δx_2, ..., Δx_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.)

An integral is an area under a curve, as described above. But we are not interested in integrals simply for the sake of studying areas of geometric objects. We study integrals largely because of their relevance to *accumulation functions.* Consider, for example, how we expressed the energy consumption of a town over a 24-hour period. The basic relation

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

could not be used directly, because power demand varies. Indirectly, though, we found that we could use this relation to build a Riemann sum for power demand p over time. This gave us an *approximation*:

$$\text{energy} \approx p(t_1)\Delta t_1 + p(t_2)\Delta t_2 + \cdots + p(t_n)\Delta t_n \quad \text{megawatt-hours.}$$

As these sums are refined (that is, more and more, narrower and narrower, rectangles are used), 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

$$\text{energy} = \int_0^{24} p(t)\,dt \quad \text{megawatt-hours} \tag{4.4.4}$$

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

In Example 4.2.1 we asked how far a car would travel in 5 hours if we knew its speed was $s(t)$ miles per hour at time t. We estimated that distance using Riemann sums for $s(t)$. Reasoning just as we did for energy, we conclude that the *exact* distance is given by the integral

$$\text{distance} = \int_0^5 s(t)\,dt \quad \text{miles.} \tag{4.4.5}$$

In other words, *distance is the integral of speed over time.*

By similar reasoning, *work is the integral of force over distance.* (See Part 1 of the Exercises for Section 4.2.) In general, whenever E is an accumulation function for a function p on an interval $[a,b]$, we have

$$\Delta E \text{ on } [a,b] = \int_a^b f(x)\,dx. \tag{4.4.6}$$

The energy integral (4.4.4) 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 t. The units for the distance integral (4.4.5) are the product of the miles per hour used to measure speed and the hours used to measure time. In general, the units for the integral in (4.4.6) are the product of the units for f and the units for x.

The process of evaluating 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.

Some integrals can be evaluated by simple geometric considerations. Let's consider a couple of integrals that can be determined in this way – and one that can't, to set the stage for the next section.

Example 4.4.1. Evaluate each of the following integrals using ideas from basic geometry, or explain why this is not possible. (If it helps, sketch the indicated function over the given interval.)

(i) $\displaystyle\int_0^5 4\,dx$ (ii) $\displaystyle\int_0^5 4x\,dx$ (iii) $\displaystyle\int_0^5 4x^2\,dx$

Solution. (i) This integral represents the area under the graph of $f(x) = 4$, a constant function of height 4, over the interval [0,5]. The region in question is just a rectangle of baselength 5 (= the length of [0,5]), and height 4 (= the height of $f(x)$). So $\int_0^5 4\,dx = 5 \times 4 = 20$.

(ii) This integral represents the area under the graph of $f(x) = 4x$, a linear function, over the interval [0,5]. The region in question is just a triangle of baselength 5 (= the length of [0,5]) and height $4 \times 5 = 20$ (= the height of $f(x)$ at $x = 5$). So $\int_0^5 4x\,dx = \frac{1}{2} \times 5 \times 20 = 50$.

(iii) This integral represents the area under the graph of $f(x) = 4x^2$, a parabola, over the interval [0,5]. The region in question has a "curved top;" this region doesn't look like any familiar shape from high school geometry. For now, we have no immediate way of determining its area.

Note that, for part (iii) of the above example, we could certainly estimate the area in question, to any desired accuracy, using Riemann sums. For example, applying the above program RIEMANN.sws to a *midpoint* approximation with $f(x) = 4x^2$, $a = 0$, $b = 5$, and $n = 5{,}000$ gives us the approximation $RS = 166.66666$ to this integral.

In the next section, we'll see how to evaluate this integral *exactly*, and we'll find that $\int_0^5 4x^2\,dx = 500/3$, which, rounded to five decimal places, equals 166.66667.

The integral of a (sometimes) negative function

Up to this point, we have been dealing with a function $y = f(x)$ that is never negative on the interval $[a, b]$. Its graph therefore lies entirely above the x-axis. What happens if f *does* take on negative values on this interval? To answer, we consider the graph below.

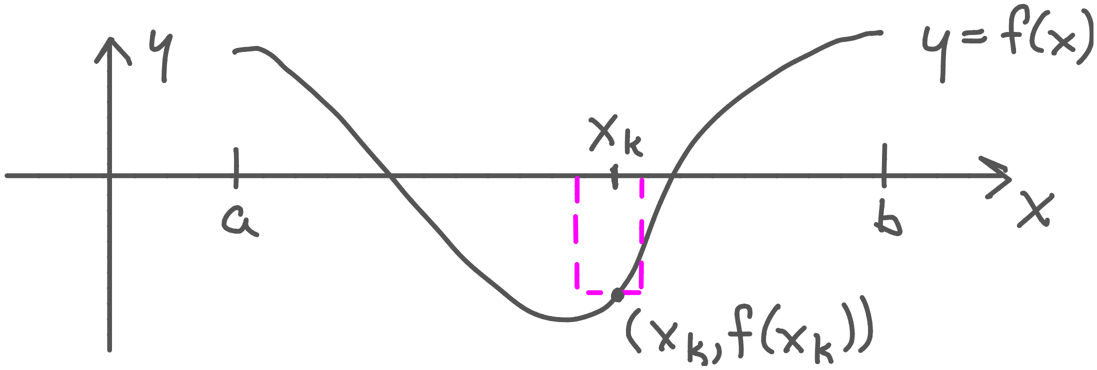


Figure 4.12. A function that's negative on part of an interval $[a,b]$, and a representative rectangle

The figure illustrates the fact that, if $f(x) < 0$ on some portion of the interval $[a,b]$, then some of the summands $f(x_k)\Delta x_k$ in the Riemann sum of (4.4.2) will, typically, be negative. And these negative contributions will impact the limit on the right-hand side of (4.4.3).

So let's imagine a series of Riemann sum rectangles spanning $[a,b]$, in Figure 4.12 above. And let's think about which rectangles lie above the x-axis, and therefore contribute a *positive* amount to the Riemann sum, and which rectangles lie below the x-axis, and therefore contribute a *negative* amount to the Riemann sum. If we then consider what happens as all of these rectangles become very narrow, we are led to the following conclusion.

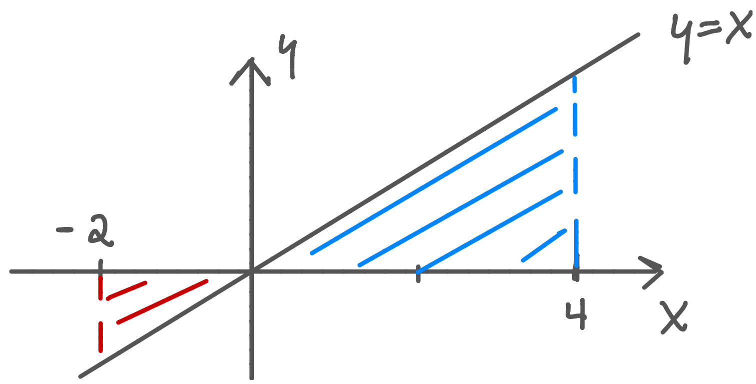
$$\int_a^b f(x)\,dx = \text{the } \textit{signed area} \text{ between}$$

the graph of $y = f(x)$ and the x-axis, on the interval $[a,b]$, meaning the sum of areas of regions above the x-axis, *minus* the sum of areas of regions below the x-axis.

Geometric interpretation of the integral, for functions that may sometimes be negative on an interval

Example 4.4.2.

(i) Consider the graph of $y = x$ over the interval $[-2, 4]$.

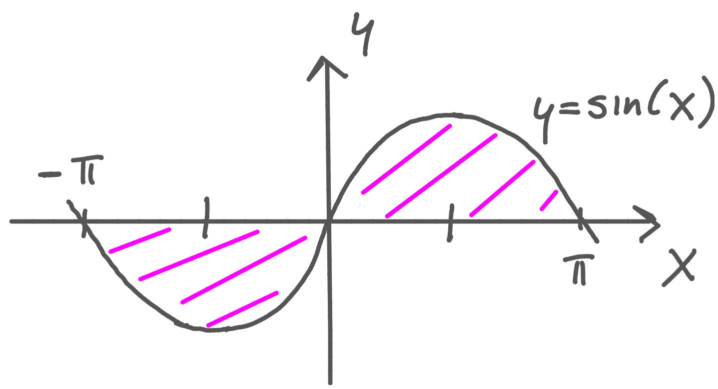


The upper region is a triangle whose area is $\frac{1}{2} \times 4 \times 4 = 8$. The lower region is another triangle; its area $\frac{1}{2} \times 2 \times 2 = 2$. Thus, the *signed area* between the graph of $y = f(x)$ and the x-axis, on the interval [−2,4], is $8 - 2 = 6$. It follows that

$$\int_{-2}^{4} x\,dx = 6.$$

You should confirm that Riemann sums for $f(x) = x$ over the interval $[-2, 4]$ converge to the value 6. See the Exercises below. (We'll evaluate this integral in another way in the next section.)

(ii) $\displaystyle\int_{-\pi}^{\pi} \sin(x)\,dx = 0$, since the areas of the two "lobes" – one above the x-axis and one below – cancel.



Again, you should confirm this result using Riemann sums. (And we'll verify it using another approach in the next section.)

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 two rules that are exactly analogous to differentiation rules:

$$\int_a^b f(x) + g(x)\,dx = \int_a^b f(x)\,dx + \int_a^b g(x)\,dx \qquad \textbf{(sum rule for integrals);}$$
$$\int_a^b c\,f(x)\,dx = c\int_a^b f(x)\,dx \qquad \textbf{(constant multiple rule for integrals).}$$

The first rule says that the signed area of a sum of two functions is the sum of the signed areas of the original two functions. The second rule says that, if you rescale a function by a factor of c in the vertical direction, then you rescale its signed area by that same factor. (And also that, if you multiply a function by a negative number c, then the new function's signed area has the opposite sign from that of the original function.)

Example 4.4.3. Use geometry, together with the sum and constant multiple rules, to compute

$$\int_{-\pi}^{\pi} (7 - 3\sin(x))\,dx.$$

Solution. By the above two rules,

$$\int_{-\pi}^{\pi} (7 - 3\sin(x))\,dx = \int_{-\pi}^{\pi} 7\,dx + \int_{-\pi}^{\pi} (-3\sin(x))\,dx = \int_{-\pi}^{\pi} 7\,dx - 3\int_{-\pi}^{\pi} \sin(x)\,dx.$$

The integral on the far right equals zero, by Example 4.4.2(ii) above. Moreover, $\int_{-\pi}^{\pi} 7\,dx = 2\pi \times 7 = 14\pi$, by the same idea as was used in Example 4.4.1(i) above. So

$$\int_{-\pi}^{\pi} (7 - 3\sin(x))\,dx = 14\pi - 3 \times 0 = 14\pi.$$

Here are two rules that have no direct analogue in differentiation.

Comparison rule for integrals: If $f(x) \leq g(x)$ for every x in the interval $[a, b]$, then

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

Juxtaposition rule for integrals: If c is a point somewhere in the interval $[a, b]$, then

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

(The sum rule tells us how integrals "add vertically;" the juxtaposition rule tells us how they "add horizontally.")

If you visualize an integral as an area, you can see why the above two rules are true. (We've already seen the juxtaposition rule in action, cf. Exercise 6(c) of Section 6.3.)

Our last rule allows us to do things like "integrate from 7 to 2," or more generally to integrate from a larger number to a smaller one. It says

$$\int_b^a f(x)\,dx = -\int_a^b f(x)\,dx \qquad \textbf{(reversal rule for integrals)}.$$

So for instance, $\int_5^0 4x\,dx = -50$, by the reversal rule and by Example 4.4.1(ii) above.

One way to think of the reversal rule is as follows: suppose $a \leq b$. If we imagine that integrating f from a to b is like traveling along $[a,b]$ from left to right, and painting the regions bounded by the graph of f, then integrating from b to a is like running that process in reverse, and "unpainting" those areas.

Still, from the point of view of mathematics and its applications, it might seem unnecessary to define integrals like $\int_5^0 4x\,dx$. As we will see, though, integrals from larger to smaller numbers *do* arise – for example, in integration by substitution, which we will consider in a later section.

Exercises

Part 1: Evaluating integrals geometrically

For these exercises, you should refer to the three examples in the section above, as well as the above integration rules (in particular, the sum rule, constant multiple rule, juxtaposition rule, and reversal rule).

1. Determine the values of the following integrals.

(a) $\int_2^{15} 3\,dx$ (b) $\int_2^{15} 3x\,dx$ (c) $\int_{15}^{2} 3x\,dx$ (d) $\int_{-\pi}^{\pi} (4\sin(x)+3)\,dx$

(e) $\int_{-\pi}^{\pi} (\sin(x)+3x)\,dx$ (f) $\int_{\pi}^{-\pi} (\sin(x)+3x)\,dx$ (g) $\int_{-4}^{9} (4-x)\,dx$

2. (a) Sketch the graph of

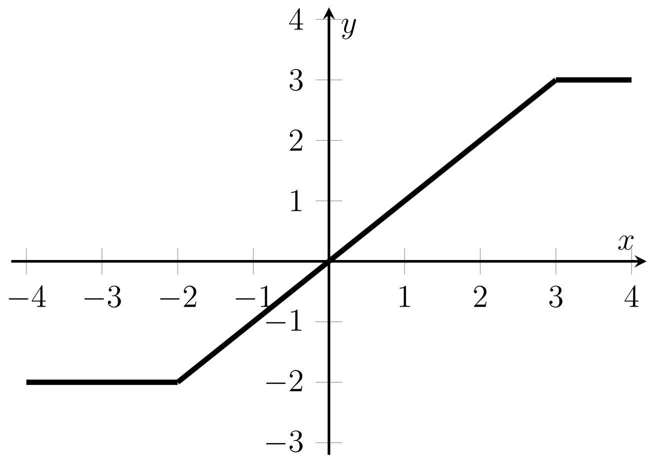
$$g(x) = \begin{cases} 7 & \text{if} \quad 1 \leq x < 5, \\ -3 & \text{if} \quad 5 \leq x \leq 10. \end{cases}$$

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

3. Repreat Exercise 2 above for

$$g(x) = \begin{cases} x-2 & \text{if} \quad 1 \leq x < 6, \\ 4 & \text{if} \quad 6 \leq x \leq 10. \end{cases}$$

4. Below is a picture of the graph of a function $y = f(x)$. Use geometry and properties of integrals to evaluate the indicated definite integrals.



(a) $\int_{-3}^{4} f(x)dx$

(b) $\int_{0}^{3} (4f(x) - 3)dx$

(c) $\int_{-4}^{1} f(x)dx + \int_{1}^{-4} f(x)dx$

5. (a) Sketch, by hand or computer, 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?

Part 2: Integrals and Riemann sums

6. Evaluate each of the following integrals, using the program RIEMANN.sws, and midpoint approximations with 2,000 rectangles. (Note: in Sage, π is entered as pi.)

(a) $\int_{-2}^{4} x\, dx$

(b) $\int_{-\pi}^{\pi} \sin(x)\, dx$

How do your results compare to those of Example 4.4.2 above?

7. Using the program RIEMANN.sws and midpoint approximations, refine Riemann sums with $n = 20, 40, 60, 80, 100, 150, 200$, and 500 rectangles, to determine the value of each of the following integrals, accurate to four decimal places. What this means is: approximate each integral with these larger and larger numbers of rectangles, until the fourth decimal place stabilizes (no longer changes).

For each integral, please provide not only your estimate (to four decimal places), but also the smallest number of rectangles you need before the fourth decimal place stabilizes. For example: if 40 rectangles give you 5.36614, 60 rectangles give you 5.36627, and the next couple of n give you $5.3662x$ (where x is any single digit), you would say "the estimate is 5.3662; it stabilizes after 60 rectangles."

(a) $\int_1^4 \sqrt{1+x^3}\,dx$ (b) $\int_4^7 \sqrt{1+x^3}\,dx$ (c) $\int_0^3 \frac{\cos(x)}{1+x^2}\,dx$ (d) $\int_0^1 \frac{1}{1+x^3}\,dx$ (e) $\int_1^2 e^{-x^2}\,dx$

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?