

We have now seen some of the most generally useful methods for discovering antiderivatives, and there are others. Unfortunately, some functions have no simple antiderivatives; in such cases if the value of a definite integral is needed it will have to be approximated. We will see two methods that work reasonably well and yet are fairly simple; in some cases more sophisticated techniques will be needed.

Of course, we already know one way to approximate an integral: if we think of the integral as computing an area, we can add up the areas of some rectangles. While this is quite simple, it is usually the case that a large number of rectangles is needed to get acceptable accuracy. A similar approach is much better: we approximate the area under a curve over a small interval as the area of a trapezoid. In figure 8.6.1 we see an area under a curve approximated by rectangles and by trapezoids; it is apparent that the trapezoids give a substantially better approximation on each subinterval.

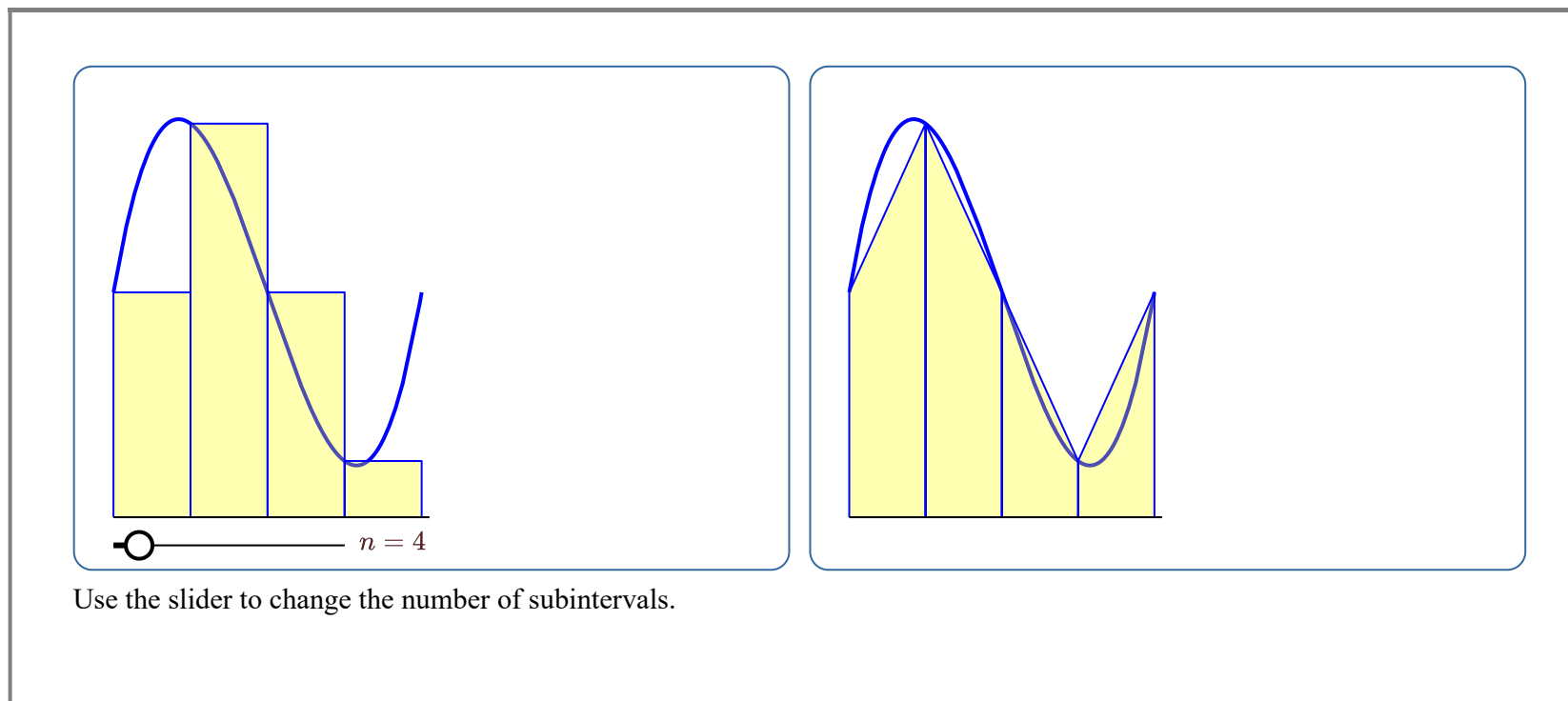


Figure 8.6.1. Approximating an area with rectangles and with trapezoids.

As with rectangles, we divide the interval into n equal subintervals of length Δx . A typical trapezoid is pictured in figure 8.6.2; it has area $\frac{f(x_i) + f(x_{i+1})}{2} \Delta x$. If we add up the areas of all trapezoids we get

$$\frac{f(x_0) + f(x_1)}{2} \Delta x + \frac{f(x_1) + f(x_2)}{2} \Delta x + \cdots + \frac{f(x_{n-1}) + f(x_n)}{2} \Delta x = \left(\frac{f(x_0)}{2} + f(x_1) + f(x_2) + \cdots + f(x_{n-1}) + \frac{f(x_n)}{2} \right) \Delta x.$$

This is usually known as the **Trapezoid Rule**. For a modest number of subintervals this is not too difficult to do with a calculator; a computer can easily do many subintervals.

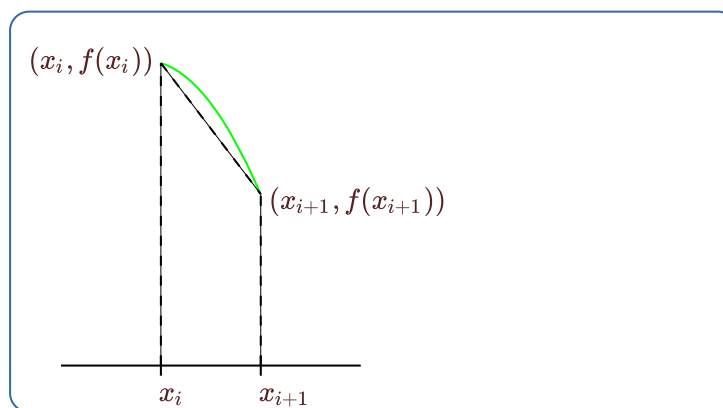


Figure 8.6.2. A single trapezoid.

In practice, an approximation is useful only if we know how accurate it is; for example, we might need a particular value accurate to three decimal places. When we compute a particular approximation to an integral, the error is the difference between the approximation and the true value of the integral. For any approximation technique, we need an **error estimate**, a value that is guaranteed to be larger than the actual error. If A is an approximation and E is the associated error estimate, then we know that the true value of the integral is between $A - E$ and $A + E$. In the case of our approximation of the integral, we want $E = E(\Delta x)$ to be a function of Δx that gets small rapidly as Δx gets small. Fortunately, for many functions, there is such an error estimate associated with the trapezoid approximation.

Theorem 8.6.1 Suppose f has a second derivative f'' everywhere on the interval $[a, b]$, and $|f''(x)| \leq M$ for all x in the interval. With $\Delta x = (b - a)/n$, an error estimate for the trapezoid approximation is

$$E(\Delta x) = \frac{b - a}{12} M (\Delta x)^2 = \frac{(b - a)^3}{12n^2} M.$$

Let's see how we can use this.

Example 8.6.2 Approximate $\int_0^1 e^{-x^2} dx$ to two decimal places. The second derivative of $f = e^{-x^2}$ is $(4x^2 - 2)e^{-x^2}$, and it is not hard to see that on $[0, 1]$, $|(4x^2 - 2)e^{-x^2}| \leq 2$. We begin by estimating the number of subintervals we are likely to need. To get two decimal places of accuracy, we will certainly need $E(\Delta x) < 0.005$ or

$$\begin{aligned}\frac{1}{12}(2)\frac{1}{n^2} &< 0.005 \\ \frac{1}{6}(200) &< n^2 \\ 5.77 &\approx \sqrt{\frac{100}{3}} < n\end{aligned}$$

With $n = 6$, the error estimate is thus $1/6^3 < 0.0047$. We compute the trapezoid approximation for six intervals:

$$\left(\frac{f(0)}{2} + f(1/6) + f(2/6) + \cdots + f(5/6) + \frac{f(1)}{2} \right) \frac{1}{6} \approx 0.74512.$$

So the true value of the integral is between $0.74512 - 0.0047 = 0.74042$ and $0.74512 + 0.0047 = 0.74982$. Unfortunately, the first rounds to 0.74 and the second rounds to 0.75, so we can't be sure of the correct value in the second decimal place; we need to pick a larger n . As it turns out, we need to go to $n = 12$ to get two bounds that both round to the same value, which turns out to be 0.75. For comparison, using 12 rectangles to approximate the area gives 0.7727, which is considerably less accurate than the approximation using six trapezoids.

In practice it generally pays to start by requiring better than the maximum possible error; for example, we might have initially required $E(\Delta x) < 0.001$, or

$$\begin{aligned}\frac{1}{12}(2)\frac{1}{n^2} &< 0.001 \\ \frac{1}{6}(1000) &< n^2 \\ 12.91 &\approx \sqrt{\frac{500}{3}} < n\end{aligned}$$

Had we immediately tried $n = 13$ this would have given us the desired answer.

The trapezoid approximation works well, especially compared to rectangles, because the tops of the trapezoids form a reasonably good approximation to the curve when Δx is fairly small. We can extend this idea: what if we try to approximate the curve more closely, by using something other than a straight line? The obvious candidate is a parabola: if we can approximate a short piece of the curve with a parabola with equation $y = ax^2 + bx + c$, we can easily compute the area under the parabola.

There are an infinite number of parabolas through any two given points, but only one through three given points. If we find a parabola through three consecutive points $(x_i, f(x_i))$, $(x_{i+1}, f(x_{i+1}))$, $(x_{i+2}, f(x_{i+2}))$ on the curve, it should be quite close to the curve over the whole interval $[x_i, x_{i+2}]$, as in figure 8.6.3. If we divide the interval $[a, b]$ into an even number of subintervals, we can then approximate the curve by a sequence of parabolas, each covering two of the subintervals. For this to be practical, we would like a simple formula for the area under one parabola, namely, the parabola through $(x_i, f(x_i))$, $(x_{i+1}, f(x_{i+1}))$, and $(x_{i+2}, f(x_{i+2}))$. That is, we should attempt to write down the parabola $y = ax^2 + bx + c$ through these points and then integrate it, and hope that the result is fairly simple. Although the algebra involved is messy, this turns out to be possible. The algebra is well within the capability of a good computer algebra system like Sage, so we will present the result without all of the algebra; you can see how to do it in this [Sage worksheet](#).

To find the parabola, we solve these three equations for a , b , and c :

$$\begin{aligned} f(x_i) &= a(x_{i+1} - \Delta x)^2 + b(x_{i+1} - \Delta x) + c \\ f(x_{i+1}) &= a(x_{i+1})^2 + b(x_{i+1}) + c \\ f(x_{i+2}) &= a(x_{i+1} + \Delta x)^2 + b(x_{i+1} + \Delta x) + c \end{aligned}$$

Not surprisingly, the solutions turn out to be quite messy. Nevertheless, Sage can easily compute and simplify the integral to get

$$\int_{x_{i+1}-\Delta x}^{x_{i+1}+\Delta x} ax^2 + bx + c \, dx = \frac{\Delta x}{3}(f(x_i) + 4f(x_{i+1}) + f(x_{i+2})).$$

Now the sum of the areas under all parabolas is

$$\begin{aligned} \frac{\Delta x}{3}(f(x_0) + 4f(x_1) + f(x_2) + f(x_2) + 4f(x_3) + f(x_4) + \cdots + f(x_{n-2}) + 4f(x_{n-1}) + f(x_n)) = \\ \frac{\Delta x}{3}(f(x_0) + 4f(x_1) + 2f(x_2) + 4f(x_3) + 2f(x_4) + \cdots + 2f(x_{n-2}) + 4f(x_{n-1}) + f(x_n)). \end{aligned}$$

This is just slightly more complicated than the formula for trapezoids; we need to remember the alternating 2 and 4 coefficients; note that n must be even for this to make sense. This approximation technique is referred to as **Simpson's Rule**.

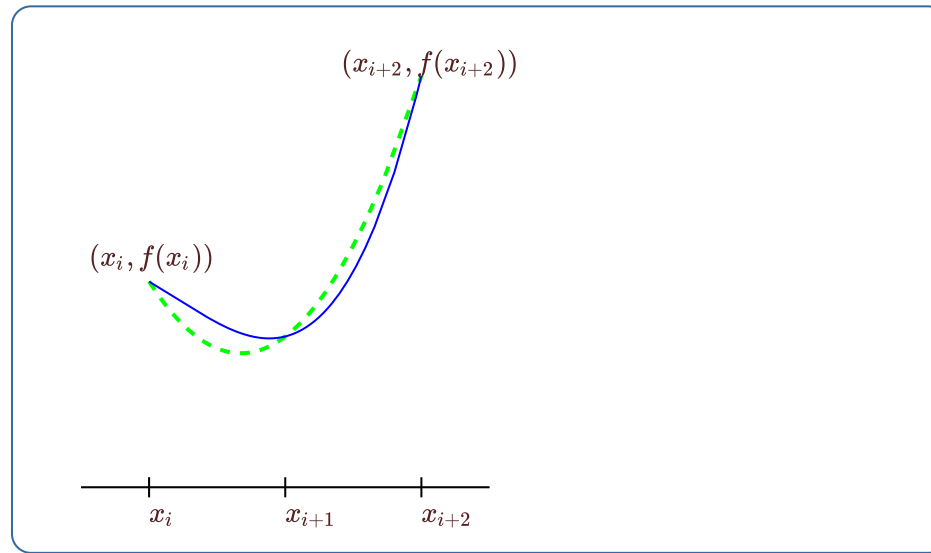


Figure 8.6.3. A parabola (dashed) approximating a curve (solid).

As with the trapezoid method, this is useful only with an error estimate:

Theorem 8.6.3 Suppose f has a fourth derivative $f^{(4)}$ everywhere on the interval $[a, b]$, and $|f^{(4)}(x)| \leq M$ for all x in the interval. With $\Delta x = (b - a)/n$, an error estimate for Simpson's approximation is

$$E(\Delta x) = \frac{b - a}{180} M (\Delta x)^4 = \frac{(b - a)^5}{180n^4} M.$$

Example 8.6.4 Let us again approximate $\int_0^1 e^{-x^2} dx$ to two decimal places. The fourth derivative of $f = e^{-x^2}$ is $(16x^2 - 48x^2 + 12)e^{-x^2}$; on $[0, 1]$ this is at most 12 in absolute value. We begin by estimating the number of subintervals we are likely to need. To get two decimal places of accuracy, we will certainly need $E(\Delta x) < 0.005$, but taking a cue from our earlier example, let's require $E(\Delta x) < 0.001$:

$$\begin{aligned} \frac{1}{180} (12) \frac{1}{n^4} &< 0.001 \\ \frac{200}{3} &< n^4 \\ 2.86 \approx \sqrt[4]{\frac{200}{3}} &< n \end{aligned}$$

So we try $n = 4$, since we need an even number of subintervals. Then the error estimate is $12/180/4^4 < 0.0003$ and the approximation is

$$(f(0) + 4f(1/4) + 2f(1/2) + 4f(3/4) + f(1))\frac{1}{3 \cdot 4} \approx 0.746855.$$

So the true value of the integral is between $0.746855 - 0.0003 = 0.746555$ and $0.746855 + 0.0003 = 0.747155$, both of which round to 0.75.