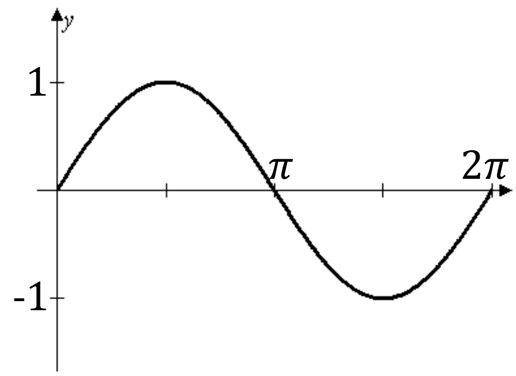
Unit 5.1: Numerical Integration

Let us first recall the geometrical interpretation of what a definite integral calculates. Suppose that f is a function that is continuous on the interval $[a, b]$. On this interval, the graph of the function along with the x-axis, bound region(s)—some above the x-axis and some below. If we integrate the function over the interval from a to b, the value of the integral will give us the area of any regions above the x-axis, minus the area of any regions below the x-axis. For example, let us consider the function $\sin x$ whose graph is given below from 0 to 2π. The graph bounds two regions each of which has area 2. When we integrate from 0 to π we obtain a value of 2 since the function is positive on that interval. However, when we integrate from π to 2π we obtain a value of -2 since the function is negative on that interval. These facts are expressed symbolically below along with two others.

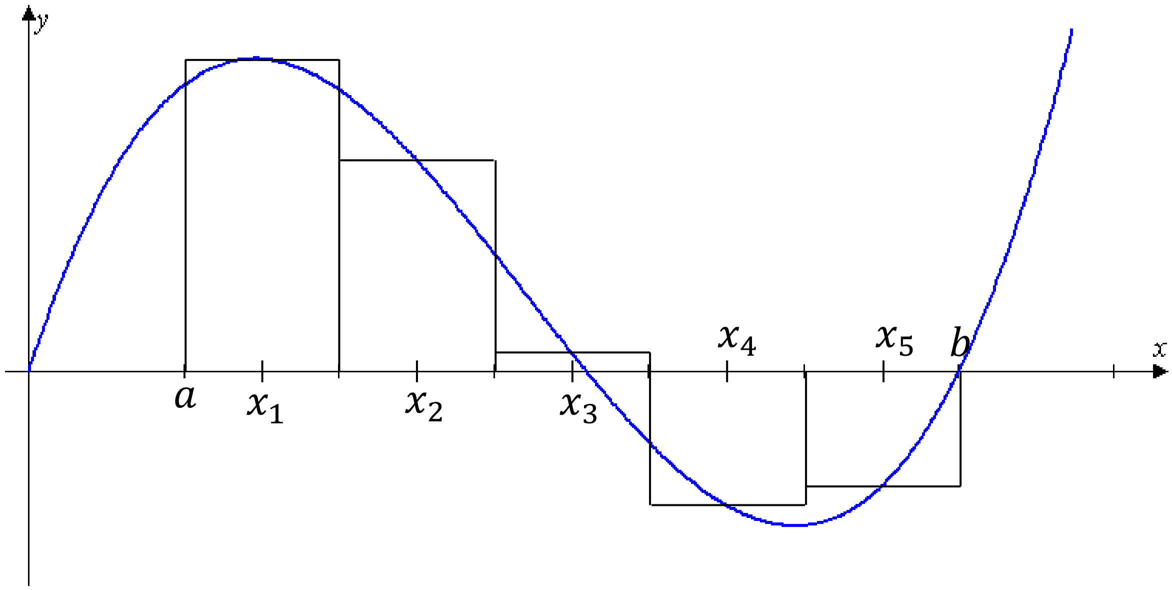


$\int_0^{\pi} \sin x \, dx = 2$ $\quad$ $\int_{\pi}^{2\pi} \sin x \, dx = -2$

$\int_0^{2\pi} \sin x \, dx = 0$ $\quad$ $\int_{\pi/2}^{2\pi} \sin x \, dx = -1$

Due to the symmetry of the regions, when we integrated from $\pi/2$ to 2π we first accumulated the integral amount of 1 as we went from $\pi/2$ to π, but then as we continued from π to 2π we accumulated an integral amount of -2, leaving us with -1 from start to finish. Each of these values could be verified using the *Fundamental Theorem of Calculus*, but the point here was to remind us of the geometrical interpretation of the integral. In order to use the *Fundamental Theorem of Calculus*, one must know the antiderivative of the integrand. This is often not the case, especially when the function is given by numerical data rather than by a formula. Thus we sometimes use numerical methods to approximate these integrals. In this section we consider the *Midpoint Rule*, the *Trapezoid Rule*, and *Simpson's Rule*.

The Midpoint Rule. Consider the integral $\int_a^b f(x)dx$, where f is continuous on $[a, b]$. We begin by partitioning the interval into n equal subintervals (n is some positive integer). The width of each of these subintervals will be denoted Δx (a small change along the x-axis). The value of Δx is given by the formula $\Delta x = \frac{b-a}{n}$. From each of these subintervals, we then choose the midpoint. The midpoint of the first subinterval will be denoted x_1, the midpoint of the second subinterval will be denoted x_2, and so on until we get to the midpoint of the n^{th} (last) subinterval which we denote x_n. Rather than referring to a specific subinterval and midpoint, we will typically speak of a generic (or representative) subinterval and midpoint and call them the i^{th} subinterval and x_i respectively. We use the value of the function at each of these midpoints as an approximation of the function on each of the corresponding intervals. When Δx is small, the function won't change much on each interval. Thus for the midpoint rule, we approximate the given function by constant functions on each of the subintervals. Instead of looking at areas under a curve, we are left looking at areas under horizontal segments, i.e. areas of rectangles. In the figure below, we show such a midpoint approximation using $n = 5$.



In approximating the integral we sum the value of $f(x_i)\Delta x$ for each of the midpoints (i.e. for $i = 1,2,3,\dots,n$). We can use Sigma notation to express such a sum as follows:

$$\int_a^b f(x)\,dx \approx \sum_{i=1}^{n} f(x_i)\Delta x = f(x_1)\Delta x + f(x_2)\Delta x + \cdots + f(x_n)\Delta x$$

For convenience, when calculating this sum, we shall factor out Δx as it is constant. We now state the **Midpoint Rule** (along with some notation).

$$\int_a^b f(x)dx \approx Mid(n) = \Delta x \cdot [f(x_1) + f(x_2) + \cdots + f(x_n)]$$

where $x_1 = a + \frac{\Delta x}{2}$, and each of the remaining midpoints can be obtained by adding Δx to the previous midpoint. For example, $x_2 = x_1 + \Delta x$. Recall, $\Delta x = \frac{b-a}{n}$.

Example 1 Approximate $\int_1^3 e^{\sin x} dx$ using $Mid(4)$.

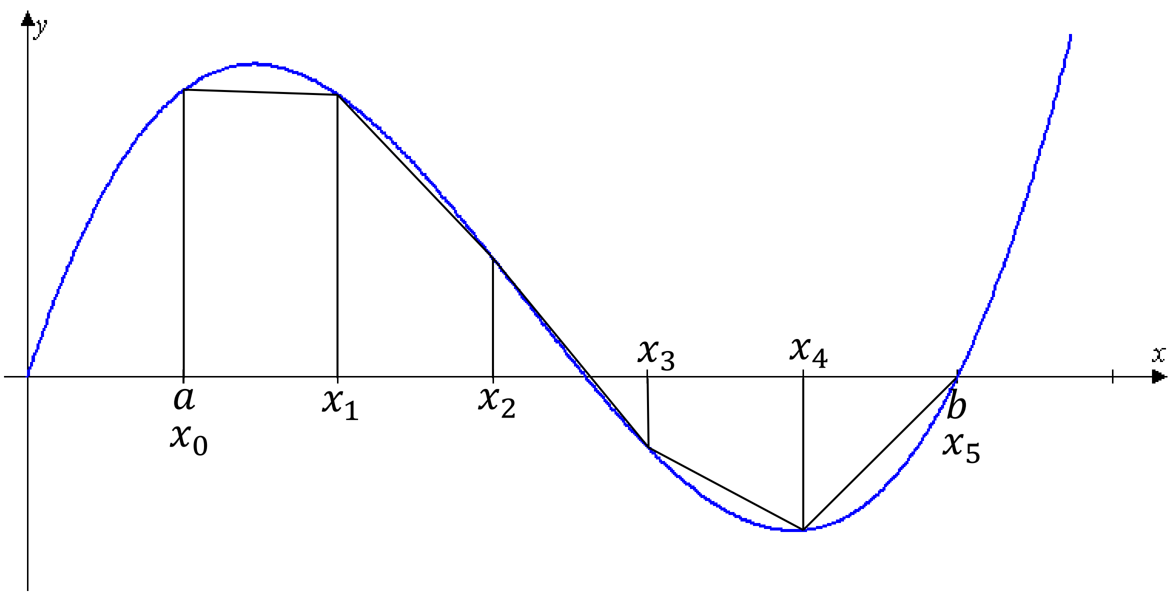
First note that we do not know the antiderivative of $e^{\sin x}$ and thus we have no hope of evaluating the integral using the *Fundamental Theorem of Calculus*. We must first calculate Δx. In this case we obtain $\Delta x = \frac{3-1}{4} = .5$ and our first midpoint is $x_1 = 1 + \frac{.5}{2} = 1.25$. Each of the remaining midpoints will be obtained by adding $.5$ to the previous midpoint. We will have four midpoints all together. We obtain

$$\int_1^3 e^{\sin x} dx \approx Mid(4) = .5\left[e^{\sin 1.25} + e^{\sin 1.75} + e^{\sin 2.25} + e^{\sin 2.75}\right] \approx \boxed{4.450}$$

The Trapezoid Rule. When we use a midpoint approximation, we are estimating the integrand with constant functions on each subinterval. The Trapezoid Rule uses linear functions (that typically increase or decrease like the integrand does) to approximate the integrand on each subinterval. With the trapezoid rule we do not use midpoints, but rather the endpoints of each of the subintervals. Let us denote these endpoints by $x_0, x_1, x_2, \ldots, x_n$. That's $(n+1)$ endpoints all together. Note: $x_0 = a$ and $x_n = b$. See the graph below where $n = 5$. At this point we will simply provide the Trapezoid Rule and refer you to the text for its derivation.

$$\int_a^b f(x)dx \approx Trap(n) = \frac{1}{2}\Delta x \cdot [f(x_0) + 2f(x_1) + 2f(x_2) + \cdots + 2f(x_{n-1}) + f(x_n)]$$

where $x_0 = a$ and each of the remaining endpoints is obtained by adding Δx to the previous endpoint. Notice the factor of 1/2 that was not present in the midpoint approximation. Also notice that all but the first and last function evaluations within the parentheses have coefficients of 2. One final and very important point is that there will be $(n+1)$ endpoints that we are substituting into our function, whereas for the midpoint rule it was just n midpoints.



<u>Example 2</u> Approximate $\int_1^3 e^{\sin x} dx$ using $Trap(4)$.

Solution: Again we have $\Delta x = \frac{3-1}{4} = .5$. We begin with the input of $x_0 = 1$ (the lower limit of integration) and continue to add $.5$ to get the remaining inputs until we have reached the input of $x_4 = 3$. At that point we should have used 5 inputs. We obtain

$$\int_1^3 e^{\sin x} dx \approx Trap(4) = \frac{1}{2}(.5)\left[e^{\sin 1} + 2e^{\sin 1.5} + 2e^{\sin 2} + 2e^{\sin 2.5} + e^{\sin 3}\right] \approx \boxed{4.375}$$

Simpson's Rule. While the Trapezoid Rule takes advantage of linear functions to account for the fact that a function is likely to change over the course of an interval, the linear approximation does not incorporate the typical curvature that our functions have. For example, approximating the position of an object with a linear function would be treating the object like it had a constant velocity. If the object were accelerating, the graph of the position function would have curvature (or concavity). To approximate the curvature, we go one step up with Simpson's Rule and use parabolic segments on each subinterval to approximate our integrand. To generate a parabolic segment, three points will be needed and thus we will use the three endpoints of two consecutive subintervals. This requires n to be even. Again, we will go straight to the formula for Simpson's Rule and refer you to the text for its derivation.

$$\int_a^b f(x)dx \approx Simp(n) = \frac{1}{3}\Delta x \cdot [f(x_0) + 4f(x_1) + 2f(x_2) + 4f(x_3) + 2f(x_4) + \cdots + 4f(x_{n-1}) + f(x_n)]$$

This rule has a similar form to the Midpoint and Trapezoid Rules, however, in this case we have a $\frac{1}{3}$ factor included and the coefficients of our function evaluations begin and end with 1, but alternate 4-2-4-2 etc…. in the "middle". Because n must be even, we will always have a coefficient of 4 on our next to last evaluation. Like in the Trapezoid Rule, we will again have $(n + 1)$ function evaluations.

<u>Example 3</u> Approximate $\int_1^3 e^{\sin x}dx$ using $Simp(4)$.

Solution: Again we have $\Delta x = \frac{3-1}{4} = .5$. We begin with the input of $x_0 = 1$ (the lower limit of integration) and continue to add $.5$ to get the remaining inputs until we have reached the input of $x_4 = 3$. At that point we should have used 5 inputs. We obtain

$$\int_1^3 e^{\sin x}dx \approx Simp(4) = \frac{1}{3}(.5)\left[e^{\sin 1} + 4e^{\sin 1.5} + 2e^{\sin 2} + 4e^{\sin 2.5} + e^{\sin 3}\right] \approx \boxed{4.427}$$

In the last three examples we approximated $\int_1^3 e^{\sin x}dx$. A summary of our results are as follows.

$$Mid(4) \approx 4.450 \qquad Trap(4) \approx 4.375 \qquad Simp(4) \approx 4.427$$

Typically (but not always), Simpson's rule gives the best estimate, however Simpson's Rule is more complicated in terms of error analysis (which we do not investigate in this course). As we are approximating these integrals, we should at least understand that our approximations are meaningless if we have no idea how close they are. We could be off by 1000 for all we know. One thing that you should know in terms of our error for a Midpoint and Trapezoid approximation is when we have an under- versus over-estimate. If the integrand is concave up (on the interval we are integrating over), a trapezoid approximation will give an overestimate, while a midpoint approximation will give an underestimate. The opposite occurs, when integrand is concave down. Convince yourself of these facts by drawing a picture.