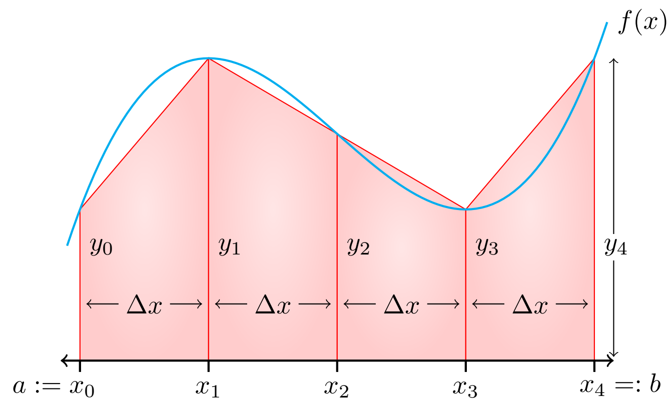
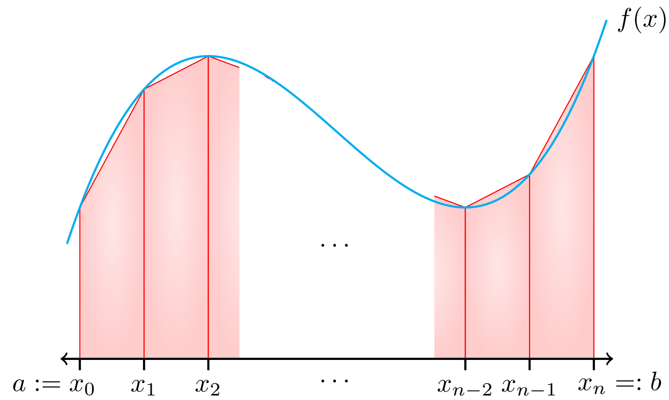
Section 8.7: Numerical Integration

What to do when there's no nice antiderivative? The antiderivatives of some functions, like $\sin(x^2)$, $1/\ln(x)$ and $\sqrt{1+x^4}$ have no elementary formulas/ When we cannot find a workable antiderivative for a function $f(x)$ that we have to integrate, we can partition the interval of integration, replace $f(x)$ by a closely fitting polynomial on each subinterval, integrate the poynomials and add the results to *approximate* the definite integral of $f(x)$. This is an example of numerical integration. There are many methods of numerical integration but we will study only two: the *Trapezium Rule* and *Simpson's Rule*.

Trapezoidal Approximations: As the name implies, the Trapezium Rule for the value of a definite integral is based on approximating the region between a curve and the x-axis with trapeziums instead of rectangles - which, if you recall, we studied when we looked at Riemann integration in Calculus I.





Assume the length of each subinterval is $\Delta x = \dfrac{b-a}{n}$. Then the area of the trapezium that lies above the x-axis in the i^{th} subinterval is $T_i = \dfrac{\delta x}{2}\left(y_{i-1} + y_i\right)$ where $y_{i-1} = f\left(x_{i-1}\right)$ and $y_i = f\left(x_i\right)$. Then the area of the under the curve and above the x-axis is approximated by the sum of the trapeziums:

$$\begin{aligned}
T &= \frac{\Delta x}{2}\left(y_0 + y_1\right) + \frac{\Delta x}{2}\left(y_1 + y_2\right) + \cdots + \frac{\Delta x}{2}\left(y_{n-1} + y_n\right) \\
&= \frac{\Delta x}{2}\left(y_0 + y_1 + y_1 + y_2 + \cdots + y_{n-2} + y_{n-1} + y_{n-1} + y_n\right) \\
&= \frac{\Delta x}{2}\left(y_0 + 2y_1 + 2y_2 + \cdots + 2y_{n-1} + y_n\right) \\
&= \frac{\Delta x}{2}\left(y_0 + y_n + 2\sum_{i=1}^{n-1} y_i\right) \\
&= \frac{\Delta x}{2}\left(f(x_0) + f(x_n) + 2\sum_{i=1}^{n-1} f\left(x_i\right)\right)
\end{aligned}$$

The Trapezium Rule: To approximate $\int_a^b f(x)\,dx$, use

$$\begin{aligned} T &= \frac{\Delta x}{2}\left(y_0 + 2y_1 + 2y_2 + \cdots + 2y_{n-1} + y_n\right) \\ &= \frac{\Delta x}{2}\left(f(x_0) + f(x_n) + 2\sum_{i=1}^{n-1} f(x_i)\right), \end{aligned}$$

where the y's are the values of f at the partition points

$$x_0 := a,\ x_1 := a + \Delta x,\ x_2 := a + 2\Delta x,\ \ldots,\ x_{n-1} := a + (n-1)\Delta x,\ x_n := a + n\Delta x = b,$$

and $\Delta x = \dfrac{b-a}{n}$.

Example 1: Use the Trapezium Rule with $n = 4$ to estimate $\int_1^2 x^2\,dx$. Compare the estimate with the exact value.

Partition the interval $[1, 2]$ into 4 subintervals:

$$\Delta x = \frac{2-1}{4} = \frac{1}{4}$$

$$x_0 = a = 1 = \frac{4}{4} \qquad x_1 = a + \Delta x = 1 + 1 \cdot \frac{1}{4} = \frac{5}{4} \qquad x_2 = a + 2\Delta x = 1 + 2 \cdot \frac{1}{4} = \frac{6}{4} \qquad x_3 = a + 3\Delta x = 1 + 3 \cdot \frac{1}{4} = \frac{7}{4} \qquad x_4 = a + 4\Delta x = 1 + 4 \cdot \frac{1}{4} = \frac{8}{4}$$

Now use these points together with the formula for the Trapezium Rule:

$$\begin{aligned} T &= \frac{\Delta x}{2}\left(y_0 + 2y_1 + 2y_2 + 2y_3 + y_4\right) \\ &= \frac{1/4}{2}\left(f\left(\frac{4}{4}\right) + 2f\left(\frac{5}{4}\right) + 2f\left(\frac{6}{4}\right) + 2f\left(\frac{7}{4}\right) + f\left(\frac{8}{4}\right)\right) \\ &= \frac{1}{8}\left(\frac{16}{16} + 2\frac{25}{16} + 2\frac{36}{16} + 2\frac{49}{16} + \frac{64}{16}\right) \\ &= \frac{1}{128}\left(16 + 50 + 72 + 98 + 64\right) \\ &= \frac{1}{128}(300) \\ &= \boxed{\frac{75}{32}} \end{aligned}$$

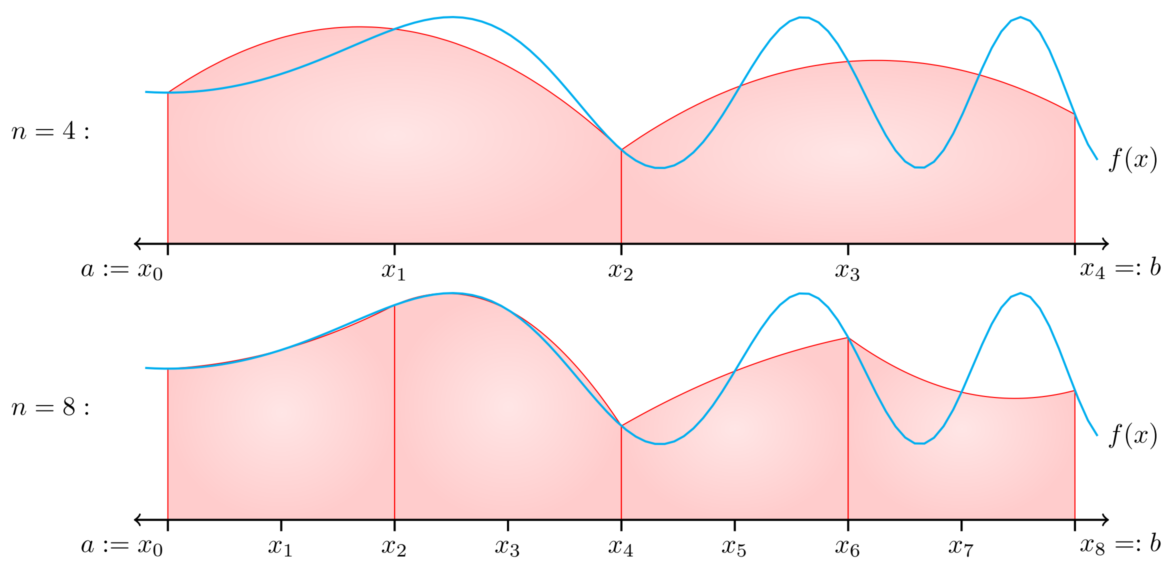
$$\begin{aligned} \int_1^2 x^2\,dx &= \frac{1}{3}x^3\Big|_1^2 \\ &= \frac{1}{3}\left(2^3 - 1^3\right) \\ &= \frac{1}{3}(8-1) \\ &= \boxed{\frac{7}{3}} \end{aligned}$$

$$\frac{75}{32} - \frac{7}{3} = \frac{225}{96} - \frac{224}{96} = \frac{1}{96}.$$

So the approximation overestimated the actual area by $\frac{1}{96}$, which is pretty good considering we only used 4 trapeziums.

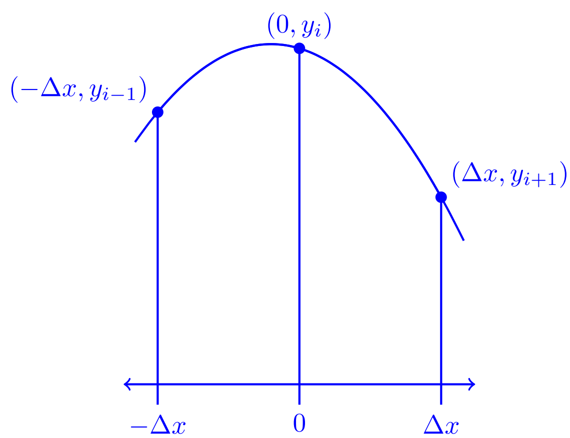
Just like when we looked at Riemann sums, using more trapeziums results in a better approximation.

Parabolic Approximations: Instead of using the straight-line segments that produced the trapeziums, we can use parabolas to approximate the definite integral of a continuous function. We partition the interval $[a, b]$ into n subintervals of equal length $\Delta x = \dfrac{b-a}{n}$ but this time we require n to be an even number. On each consecutive pair of intervals we approximate the curve $y = f(x) \geq 0$ by a parabola. A typical parabola passed through three consecutive points: (x_{i-1}, y_{i-1}), (x_i, y_i) and (x_{i+1}, y_{i+1}) on the curve.



So how do we compute the area under each parabola $y = Ax^2 + Bx + C$? By translating we can assume that the centre point of our parabola is at $x_i = 0$

The area under the parabola and above the x-axis is given by



$$
\begin{aligned}
S_i &= \int_{-\Delta x}^{\Delta x} Ax^2 + Bx + C\,dx \\
&= \frac{Ax^3}{3} + \frac{Bx^2}{2} + Cx\Bigg|_{-\Delta x}^{\Delta x} \\
&= \frac{A(\Delta x)^3}{3} + \frac{B(\Delta x)^2}{2} + C(\Delta x) - \left[\frac{A(-\Delta x)^3}{3} + \frac{B(-\Delta x)^2}{2} + C(-\Delta x)\right] \\
&= \frac{2A\Delta x^3}{3} + 2C\Delta x \\
&= \frac{\Delta x}{3}\left(2A\Delta x + 6C\right)
\end{aligned}
$$

$$
\left.\begin{aligned}
y_{i-1} &= A\Delta x^2 - B\Delta x + C \\
y_i &= C \\
y_{i+1} &= A\Delta x^2 + B\Delta x + C
\end{aligned}\right\} \implies y_{i-1} + 4y_i + y_{i+1} = \left(A\Delta x^2 - B\Delta x + C\right) + 4C + \left(A\Delta x^2 + B\Delta x + C\right) = 2A\Delta x + 6C
$$

$$
\implies \boxed{S_i = \frac{\Delta x}{3}\left(y_{i-1} + 4y_i + y_{i+1}\right)}
$$

So if we sum up the areas under all of the parabolas, we obtain our approximation.

Simpson's Rule: To approximate $\int_a^b f(x)\,dx$, use

$$\begin{aligned} S &= \frac{\Delta x}{3}\left(y_0 + 4y_1 + 2y_2 + 4y_3 + \cdots + 2y_{n-2} + 4y_{n-1} + y_n\right) \\ &= \frac{\Delta x}{3}\left(f(x_0) + f(x_n) + 2\left(\sum_{i=1}^{\frac{n-1}{2}} f(x_{2i-1}) + 2f(x_{2i})\right)\right), \end{aligned}$$

where the y's are the values of f at the partition points

$$x_0 := a,\ x_1 := a + \Delta x,\ x_2 := a + 2\Delta x,\ \ldots,\ x_{n-1} := a + (n-1)\Delta x,\ x_n := a + n\Delta x = b,$$

and $\Delta x = \dfrac{b-a}{n}$ with n an *even* number.

Example 2: Use the Simpson's Rule with $n = 4$ to approximate $\int_0^2 5x^4\,dx$. Compare the estimate with the exact value.

Partition the interval $[1, 2]$ into 4 subintervals:

$$\Delta x = \frac{2-0}{4} = \frac{1}{2}$$

$$x_0 = a = 0 = \frac{0}{2}$$

$$x_1 = a + \Delta x = 0 + 1 \cdot \frac{1}{2} = \frac{1}{2}$$

$$x_2 = a + 2\Delta x = 0 + 2 \cdot \frac{1}{2} = \frac{2}{2}$$

$$x_3 = a + 3\Delta x = 0 + 3 \cdot \frac{1}{2} = \frac{3}{2}$$

$$x_4 = a + 4\Delta x = 0 + 4 \cdot \frac{1}{2} = \frac{4}{2}$$

Now use these points together with the formula for Simpson's Rule:

$$\begin{aligned} S &= \frac{\Delta x}{3}\left(y_0 + 4y_1 + 2y_2 + 4y_3 + y_4\right) \\ &= \frac{1/2}{3}\left(f\left(\frac{0}{2}\right) + 4f\left(\frac{1}{2}\right) + 2f\left(\frac{2}{2}\right) + 4f\left(\frac{3}{2}\right) + f\left(\frac{4}{2}\right)\right) \\ &= \frac{1}{6}\left(5\frac{0}{16} + 4 \cdot 5\frac{1}{16} + 2 \cdot 5\frac{16}{16} + 4 \cdot 5\frac{81}{16} + 5\frac{128}{16}\right) \\ &= \frac{5}{96}(0 + 4 + 32 + 324 + 256) \\ &= \frac{5}{96}(616) \\ &= \boxed{\frac{385}{12}} \end{aligned}$$

$$\begin{aligned} \int_0^2 5x^4\,dx &= x^5\Big|_0^2 \\ &= 2^5 - 0^5 \\ &= 32 - 0 \\ &= \boxed{32} \end{aligned}$$

$$\frac{385}{12} - 32 = \frac{385}{12} - \frac{384}{12} = \frac{1}{12}.$$

So the approximation overestimated the actual area by $\frac{1}{12}$, which is pretty good considering we only used 2 parabolas.

Just like Riemann sums and the Trapezium rule, using more parabolas results in a better approximation. In fact, of the three rules Simpson's Rule gives the best approximation. This can be seen by looking at the *error estimates*.

Error Estimates in the Trapezium and Simpson's Rules If $f''(x)$ is continuous and M is any upper bound for the values of $|f''(x)|$ on $[a, b]$, then the error E_T in the Trapezium Rule for approximating the definite integral of $f(x)$ over the interval $[a, b]$ using n trapeziums satisfies the inequality

$$|E_T| \leq \frac{M(b-a)^3}{12n^2}.$$

If $f^{(4)}(x)$ is continuous and M is any upper bound for the values of $|f^{(4)}(x)|$ on $[a, b]$, then the error E_S in Simpson's Rule for approximating the definite integral of $f(x)$ over the interval $[a, b]$ using $\frac{n}{2}$ parabolas satisfies the inequality

$$|E_S| \leq \frac{M(b-a)^5}{180n^4}.$$

Example 3: Find an upper bound for the error in estimating $\int_0^2 5x^4\,dx$ using Simpson's Rule with $n = 4$. What value of n should we pick so that the error is within 0.001 of the true value?

First we differentiate $f(x)$ 4 times and check that it is continuous on the interval $[0, 2]$.

$$\begin{aligned} f(x) &= 5x^4 \\ f'(x) &= 20x^3 \\ f''(x) &= 60x^2 \\ f'''(x) &= 120x \\ f^{(4)}(x) &= 120 \end{aligned}$$

This is a constant function, so it *is* continuous on our interval. Further

$$|f^{(4)}(x)| = 120 \leq 120 \text{ for all } x \in [0, 2].$$

Thus $M = 120$ works as a bound. So, with $n = 4$, the error is bounded by:

$$|E_S| \leq \frac{M(b-a)^5}{180n^4} = \frac{120(2-0)^5}{180(4)^4} = \frac{120 \cdot 2^5}{180 \cdot 2^8} = \frac{1}{3 \cdot 2^2} = \frac{1}{12}.$$

To achieve an approximation with $|E_S| \leq 0.001$, we again find a bound for M but this time we solve the inequality for n.

$$\frac{M(b-a)^5}{180n^4} = \frac{120(2-0)^5}{180n^4} = \frac{2^6}{3n^4} \leq 0.001$$

$$\begin{aligned} \implies \quad & \frac{2^6}{3} \leq \frac{1}{1000}n^4 \\ \implies \quad & \frac{2^6 \cdot 1000}{3} \leq n^4 \\ \implies \quad & \frac{2^8 \cdot 2 \cdot 5^3}{3} \leq n^4 \\ \implies \quad & 4\sqrt[4]{\frac{2 \cdot 5^3}{3}} \leq n \end{aligned}$$

So setting $n \geq 4\sqrt[4]{\frac{2\cdot 5^3}{3}} \approx 12.086$ would ensure an approximation of the desired accuracy.