Math 340 • Fall 2026 • Victor Matveev

5.2 Numerical Integration

Even many elementary functions often do not have closed-form anti-derivatives, and can only be integrated numerically. Here are just a couple important examples:

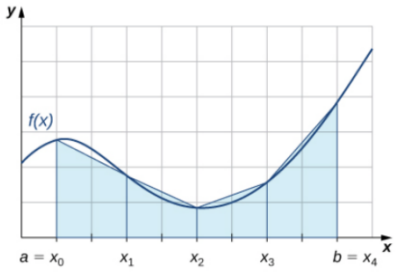
Error function: $\int_a^b \exp\left(-x^2\right)dx$

Fresnel integral: $\int_a^b \sin\left(x^2\right)dx$

Another example: $\int_a^b \exp\left(\sin(x)\right)dx$

Fortunately, we can find the integral of any piece-wise continuous function quite accurately using numerical methods.

One of the simplest methods learned in Calculus I is the **Trapezoidal Rule**, which involves approximating the area under the function curve with the area of trapezoids formed by the function values at the partition points, as shown in the figure to the right:



For a single subinterval in this figure, the integral is approximated by the area of **one** trapezoid (equal to base length times average height), which we denote $\boldsymbol{T_1}$:

$$\int_{x_0}^{x_1} f(x)dx \approx T_1 = \frac{h}{2}\left[f(x_0)+f(x_1)\right], \quad \text{where } h = x_1 - x_0.$$

The figure makes it clear that this integration method is exact for a linear function, but is only approximate otherwise.

To analyze the error of this approximation rigorously, we will check how well the method works for monomials of increasing order. Without loss of generality, let's consider the interval of integration to be $[0, h]$:

$f(x)$	Approximation $T_1[f] \equiv \frac{h}{2}\left[f(0)+f(h)\right]$	Exact integral $I[f] \equiv \int_0^h f(x)dx$	Error $T_1 - I[f]$
1	$\frac{h}{2}[1+1] = h$	$\int_0^h dx = h$	0
x	$\frac{h}{2}[0+h] = \frac{h^2}{2}$	$\int_0^h x\,dx = \frac{h^2}{2}$	0
x^2	$\frac{h}{2}\left[0+h^2\right] = \frac{h^3}{2}$	$\int_0^h x^2 dx = \frac{h^3}{3}$	$\frac{h^3}{2} - \frac{h^3}{3} = \frac{h^3}{6}$

For any function that has a continuous 2nd order derivative, the linearity of the integral means that the error is linear as well, i.e. **the error of a linear combination is a linear combination of errors**:

$$Error\left[f(x)\right] = Error\left[f(0)\cdot 1 + f'(0)\cdot x + \frac{f''(c_x)}{2}\cdot x^2\right]$$

$$= f(0)\cdot \underbrace{Error[1]}_{0} + f'(0)\cdot \underbrace{Error[x]}_{0} + \frac{f''(C)}{2}\cdot \underbrace{Error\left[x^2\right]}_{h^3/6} = \boxed{\frac{f''(C)}{12}h^3} \quad \text{where } C \in (0,h)$$

The only non-obvious part in this error derivation is the fact the $f''(c_x)$ values appearing in the Taylor remainder term average out to some overall mean value $f''(C)$ when integrated over the interval $[0, h]$, with $C \in (0, h)$.

Note: this averaging corresponds to the mean value theorem for integrals: $\exists c \in (a,b): \int_a^b f(x)h(x)dx = f(c)\int_a^b h(x)dx.$

EXAMPLE: Use trapezoidal rule to approximate $\int_1^2 \frac{dx}{x} = \ln(2) \approx \mathbf{0.6931472}.$.

Solution: $T_1 = \frac{h}{2}\left[f(1)+f(2)\right] = \frac{1}{2}\left[1+\frac{1}{2}\right] = \frac{3}{4} = \mathbf{0.75}.$

Actual error: $T_1 - \ln(2) = \mathbf{0.057}.$

Theoretical error bound: $\left|\frac{f''(C)}{12}h^3\right| = \left|\frac{2/C^3}{12}\right| \le \frac{1}{6} = \mathbf{0.1666}..$

Thus, actual error is well within the theoretical error bound.

Composite Trapezoidal Rule

To increase integration accuracy, we may subdivide the integration interval into many partitions, as shown in the figure on the previous page, and then add up the resulting trapezoidal areas. This results in the **composite trapezoidal rule**:

$$\int_{a=x_0}^{b=x_n} f(x)dx \approx T_n = \frac{h}{2}\left[f(x_0)+f(x_1)\right] + \frac{h}{2}\left[f(x_1)+f(x_2)\right] + \ldots + \frac{h}{2}\left[f(x_{n-2})+f(x_{n-1})\right] + \frac{h}{2}\left[f(x_{n-1})+f(x_n)\right]$$

$$\boxed{T_n = \frac{h}{2}\left[f(x_0)+2f(x_1)+2f(x_2)+\ldots+2f(x_{n-1})+f(x_n)\right]} \quad \text{where } h = x_{n+1} - x_n.$$

Of course, the errors add up as well, so the error of this approximation is

$$Error\left[T_n[f]\right] = \sum_{k=1}^{n} \frac{f''(C_k)}{12}h^3 = \frac{nf''(C)}{12}h^3 = \boxed{\frac{f''(C)(b-a)}{12}h^2} \quad \text{where } C \in (a,b),\ n \cdot h = (b-a)$$

Here we used averaging implied by the intermediate value theorem: $\exists C \in (a,b): \sum_{k=1}^{n} f''(C_k) = nf''(C)$

This expression reveals the **2nd order accuracy of the composite Trapezoidal rule**.

Another useful concept in addition to order of accuracy is the **degree of precision** of a method, which is the highest degree of polynomial for which the method is exact. For the trapezoidal rule, **the degree or precision = 1**, since the error is zero for linear function since $f''(x) = 0$ in this case.

EXAMPLE: Use **composite** trapezoidal rule with $\boldsymbol{n = 2}$ subintervals to approximate $\int_1^2 \frac{dx}{x} = \ln(2) \approx \mathbf{0.6931472}.$

Solution: $\mathbf{T_2} = \frac{h}{2}\left[f(1)+2f\left(\frac{3}{2}\right)+f(2)\right] = \frac{1}{4}\left[1+2\cdot\frac{2}{3}+\frac{1}{2}\right] = \frac{17}{24} \approx \mathbf{0.7083}.$

Actual error: $T_2 - \ln(2) = \mathbf{0.0152}.$

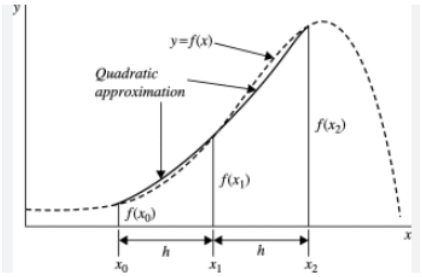
Theoretical error bound: $\left|\frac{f''(C)}{12}(b-a)h^2\right| = \left|\frac{2/C^3}{12}\left(\frac{1}{2}\right)^2\right| \le \frac{1}{24} \approx \mathbf{0.0417}.$

Thus, once again the actual error is well within the theoretical error bound.

Simpson's Rule.

Trapezoidal rule is equivalent to replacing the integrand $f(x)$ with its 2-point polynomial interpolation, which is a line. If we use **3 points** for interpolation, we are replacing the integrand with its quadratic interpolating polynomial before integrating. This results in a rule known as the Simpson's Rule, named after Thomas Simpson (1710-1761).

For a single partition in this figure, the resulting area under the parabola will be denoted $\boldsymbol{S_2}$, and, as can be shown by integrating a quadratic 3-point polynomial interpolant in Lagrange form, equals



$$\int_{x_0}^{x_2} f(x)dx \;\approx\; S_2 = \frac{h}{3}\left[f(x_0)+4f(x_1)+f(x_2)\right], \quad \text{where } h = x_{k+1}-x_k.$$

The explanation above makes it clear that this integration method is exact for a quadratic polynomial. As we will see, its degree of precision is even better than that!

To analyze the error of Simpsons' rule, we will once again check how well the method works for monomials of increasing order. Without loss of generality, let's consider the interval of integration of $[-h,h]$, namely $x_0 = -h\,, x_1 = 0,\ x_2 = h$:

$f(x)$	Approximation $S_2[f] \equiv \frac{h}{3}\left[f(-h)+4f(0)+f(h)\right]$	Exact integral $I[f] \equiv \int_{-h}^{h} f(x)dx$	Error $Error[f] \equiv S_2 - I[f]$
1	$\frac{h}{3}[1+4+1] = 2h$	$\int_{-h}^{h} dx = 2h$	0
x	$\frac{h}{3}[-h+0+h] = 0$	$\int_{-h}^{h} x\,dx = 0$	0
x^2	$\frac{h}{3}\left[(-h)^2+0+h^2\right] = \frac{2h^3}{3}$	$\int_{-h}^{h} x^2 dx = \frac{2h^3}{3}$	0
x^3	$\frac{h}{3}\left[(-h)^3+0+h^3\right] = 0$	$\int_{-h}^{h} x^3\,dx = 0$	0
x^4	$\frac{h}{3}\left[(-h)^4+0+h^4\right] = \frac{2h^5}{3}$	$\int_{-h}^{h} x^4 dx = \frac{2h^5}{5}$	$2h^5\left(\frac{1}{3}-\frac{1}{5}\right) = \frac{4h^5}{15}$

For any function that has a continuous 4th order derivative, the linearity of the integral and its numerical approximation allows us to write the error as

$$Error\left[f(x)\right] = Error\left[f(0)\cdot 1 + f'(0)\cdot x + \frac{f''(0)}{2!}\cdot x^2 + \frac{f'''(0)}{3!}\cdot x^3 + \frac{f^{(4)}(c_x)}{4!}\cdot x^4\right]$$

$$= \underbrace{Error\left[f(0)\cdot 1 + f'(0)\cdot x + \frac{f''(0)}{2!}\cdot x^2 + \frac{f'''(0)}{3!}\cdot x^3\right]}_{=0} + \frac{f^{(4)}(C)}{4!}\cdot \underbrace{Error\left[x^4\right]}_{4h^5/15} = \boxed{\frac{f^{(4)}(C)}{90}h^5} \quad \text{where } C \in (-h,\,h)$$

Side note: once again, the $f^{(4)}(c_x)$ values appearing in the Taylor remainder term average out to some overall mean value $f^{(4)}(C)$ when integrated over the interval $[-h,h]$, with $C \in (-h,h)$.

EXAMPLE: Use elementary Simpson's rule to approximate $\int_1^2 \frac{dx}{x} = \ln(2) \approx \mathbf{0.6931472}$.

Solution: $\mathbf{S_2} = \frac{h}{2}\left[f(1)+4f\left(\frac{3}{2}\right)+f(2)\right] = \frac{1}{4}\left[1+4\cdot\frac{2}{3}+\frac{1}{2}\right] = \frac{25}{36} \approx \mathbf{0.69444}$.

Actual error: $S_2 - \ln(2) = \mathbf{0.0013}$.

Theoretical error bound: $\left|\frac{f^{(4)}(C)}{90}h^5\right| = \left|\frac{24/C^5}{90}\left(\frac{1}{2}\right)^5\right| \le \left|\frac{24}{90\cdot 32}\right| = \left|\frac{1}{30\cdot 4}\right| = \frac{1}{120} \approx \mathbf{0.008333}...$

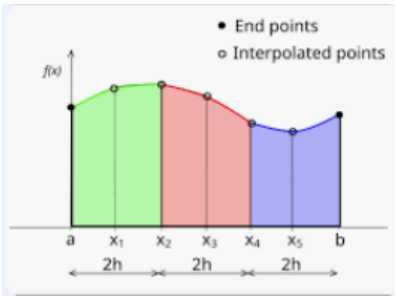
Once again, the actual error is well within the theoretical error bound.

Note that the actual **error is 10 times smaller** than the error of the Trepezoidal approximation with 2 partitions, even though we use the **same three points** to estimate the integral!

We also obtained a curious rational approximation: $\ln(2) \approx \left(\frac{5}{6}\right)^2$ ☺

Composite Simpson's Rule

To increase integration accuracy, we may subdivide the integration interval into an even number of partitions, as shown in the figure to the right, and then add up the resulting areas for each interval pair. This results in the **composite Simpson's rule**:



$$\int_{a=x_0}^{b=x_n} f(x)dx \approx S_n = \frac{h}{3}\left[f(x_0)+4f(x_1)+f(x_2)\right]+\frac{h}{3}\left[f(x_2)+4f(x_3)+f(x_4)\right]+\ldots+\frac{h}{3}\left[f(x_{n-2})+4f(x_{n-1})+f(x_n)\right]$$

$$\boxed{S_n = \frac{h}{3}\left[f(x_0)+4f(x_1)+2f(x_2)+4f(x_3)+\ldots+2f(x_{n-2})+4f(x_{n-1})+f(x_n)\right]} \quad \text{where } h = x_{n+1}-x_n.$$

Since there are $(n/2)$ partition pairs given an even number of partitions n, the error of this approximation is

$$Error\left[S_n[f]\right] = \sum_{k=1}^{n/2}\frac{f^{(4)}(C_k)}{90}h^5 = \left(\frac{n}{2}\right)\frac{f^{(4)}(C)}{90}h^5 = \boxed{\frac{f^{(4)}(C)(b-a)}{180}h^4} \quad \text{where } C\in(a,b),\ n\cdot h=(b-a)$$

Here we used again the averaging implied by the intermediate value theorem: $\exists C\in(a,b):\ \sum_{k=1}^{n/2} f^{(4)}(C_k) = \frac{n}{2}f^{(4)}(C)$

This expression reveals the **4th order accuracy of the composite Simpson's rule**.

The **degree of precision of the Simpson's rule equals 3**, since we showed it to be exact for 3rd degree polynomials.

Richardson Extrapolation

Richardson extrapolation is a clever way of combining a numerical approximation of certain accuracy with a less accurate approximation to obtain a result that is more accurate than either of the two approximations.

Let's consider for instance the Trapezoidal rule, which we have shown to have order of accuracy $\boldsymbol{n = 2}$:

$$T_n = \int_a^b f(x)dx + \frac{f''(c)(b-a)}{12}h^2$$

Denoting the exact value of the integral as $\boldsymbol{Q}$, and the error coefficient as $\boldsymbol{K = \frac{f''(c)(b-a)}{12}}$, we can write the above as

$$T_n = Q + Kh^2$$

Assuming n is even, we can then consider the error arising from using a less accurate approximation with only half the partition points, in which case subpartition interval *h* is twice as high:

$$T_{n/2} = Q + K(2h)^2 = Q + 4Kh^2$$

Thus, we obtain a simple linear system with 2 unknowns:

$$\begin{cases} T_n = Q + Kh^2 \\ T_{n/2} = Q + 4Kh^2 \end{cases}$$

Note that we can easily eliminate the error term using linear combination $4T_n - T_{n/2} \approx 3Q$:

$$\boldsymbol{Q \approx \frac{4T_n - T_{n/2}}{3}}$$

This elegant trick allows us to improve the accuracy of any approximation by combining it with a less accurate approximation! 😊

NOTE: Since K is only approximately independent of h, the value Q is not completely error-free. However, this rigorously eliminates terms of order h^2 in the error, and any remaining error in Q is of higher order, h^4.

When $\boldsymbol{n = 1}$, this equation combines the elementary Trapezoidal rule with the 3-point composite Trapezoidal rule; let's see what we obtain in this case (considering that the partition size *h* is twice as large for T_1 compared to T_2):

$$\begin{cases} \mathbf{T_2} = \frac{h}{2}\left[f(x_0) + 2f(x_1) + f(x_2)\right] = h\left[\frac{1}{2}f(x_0) + f(x_1) + \frac{1}{2}f(x_2)\right] \\ \mathbf{T_1} = \frac{2h}{2}\left[f(x_0) + f(x_2)\right] = h\left[f(x_0) + f(x_2)\right] \end{cases}$$

$$Q = \frac{1}{3}(4T_2 - T_1) = \frac{h}{3}\cdot\left[2f(x_0) + 4f(x_1) + 2f(x_2) - f(x_0) - f(x_2)\right] = \frac{h}{3}\cdot\left[f(x_0) + 4f(x_1) + f(x_2)\right] = S_2$$

Thus, **Richardson Extrapolation applied to Trapezoidal Rule transforms it to Simpson's Rule!** 😊

EXAMPLE: Examine Richardson Extrapolation of T_1 and T_2 Trapezoidal Approximations to the integral

$$\int_1^2 \frac{dx}{x} = \ln(2) \approx \mathbf{0.6931472}.$$

Solution: We found on previous pages that $\boldsymbol{T_1 = \frac{3}{4}}$ and $\boldsymbol{T_2 = \frac{17}{24}}$. Thus, a more accurate extrapolated value is

$$\boldsymbol{Q = \frac{1}{3}(4T_2 - T_1) = \frac{1}{3}\left(4\cdot\frac{17}{24} - \frac{3}{4}\right) = \frac{1}{3}\left(\frac{17}{6} - \frac{3}{4}\right) = \frac{1}{6}\left(\frac{17}{3} - \frac{3}{2}\right) = \frac{25}{36}}$$

Note that we obtained the Simpson's Rule value, as expected!

Richardson Extrapolation: General Case, and Boole's Rule

In general, if some expression $\boldsymbol{F(h)}$ approximates a quantity $\boldsymbol{Q}$ with order of accuracy $\boldsymbol{n}$, the same logic leads to:

$$\begin{cases} F(h) = Q + Kh^n \\ F(2h) = Q + K(2h)^n = Q + 2^n\left(Kh^n\right) \end{cases}$$

Thus, we can eliminate the error term using the following linear combination of a more accurate and a less accurate approximation:

$$Q = \frac{2^n F(h) - F(2h)}{2^n - 1}$$

Let's apply this to Simpson's Rule, which we found to have **order of accuracy of $\boldsymbol{n = 4}$**:

$$\begin{cases} F(h) = S_4 = \frac{h}{3}\left[f(x_0) + 4f(x_1) + 2f(x_2) + 4f(x_3) + f(x_4)\right], \\ F(2h) = S_2 = \frac{2h}{3}\left[f(x_0) + 4f(x_2) + f(x_4)\right] \\ \qquad\quad = \frac{h}{3}\left[2f(x_0) + 8f(x_2) + 2f(x_4)\right]. \end{cases}$$

$$\begin{aligned} Q = \frac{2^4 S_4 - S_2}{2^4 - 1} &= \frac{1}{15}(16S_4 - S_2) \\ &= \frac{h}{45}\cdot\left[16f(x_0) + 64f(x_1) + 32f(x_2) + 64f(x_3) + 16f(x_4) - 2f(x_0) - 8f(x_2) - 2f(x_4)\right] \\ &= \frac{h}{45}\cdot\left[14f(x_0) + 64f(x_1) + 24f(x_2) + 64f(x_3) + 14f(x_4)\right] \\ &= \frac{2h}{45}\cdot\left[7f(x_0) + 32f(x_1) + 12f(x_2) + 32f(x_3) + 7f(x_4)\right] \end{aligned}$$

We obtained the so-called **Boole's Rule**, which is much more accurate than Simpson's rule!

The error of the basic Boole's Rule is $\frac{\boldsymbol{8f^{(6)}(c)}}{\boldsymbol{945}}\boldsymbol{h^7}$, so the error of composite rule is $\frac{n}{4}\cdot\frac{8f^{(6)}(C)}{945}h^7 = \frac{\boldsymbol{2(b-a)f^{(6)}(C)}}{\boldsymbol{945}}\boldsymbol{h^6}$

Continued application of Richardson Extrapolation to approximants of increasing order of accuracy is the essence of **Romberg Integration** explained in class.

EXAMPLE: Use elementary Boole's rule to approximate $\int_1^2 \frac{dx}{x} = \ln(2) \approx \mathbf{0.6931472}$.

Solution:

$$B_4 = \frac{2h}{45}\left[7f(1) + 32f\left(\frac{5}{4}\right) + 12f\left(\frac{3}{2}\right) + 32f\left(\frac{7}{4}\right) + 7f(2)\right] = \frac{1}{90}\left[7 + \frac{128}{5} + \frac{24}{3} + \frac{128}{7} + \frac{7}{2}\right] = \mathbf{0.6931746}\ldots$$

Actual error: $B_4 - \ln(2) \approx \mathbf{2.7\times10^{-5}}$

Theoretical error bound: $\left|\frac{8f^{(6)}(C)}{945}h^7\right| = \left|\frac{8\cdot 6!/C^7}{945}\left(\frac{1}{4}\right)^7\right| \le \left|\frac{1}{21\cdot 128}\right| \approx \mathbf{3.7\times10^{-4}}$

Summary: Newton-Cotes Formulas

Integration methods that we considered so far all involve equipartition of the integration interval – such methods are collectively called Newton-Cotes methods, and the ones we reviewed are summed up below:

Trapezoidal: $T_1 = \frac{h}{2}\left[f(x_0) + f(x_1)\right]$, $\text{Error} = \frac{f''(c)}{12}h^3$, $\text{Composite error} = \frac{(b-a)f''(c)}{12}h^2$

Simpson's: $S_2 = \frac{h}{3}\left[f(x_0) + 4f(x_1) + f(x_2)\right]$, $\text{Error} = \frac{f^{(4)}(c)}{90}h^5$, $\text{Composite error} = \frac{(b-a)f^{(4)}(c)}{180}h^4$

Boole's: $B_4 = \frac{2h}{45}\left[7f(x_0) + 32f(x_1) + 12f(x_2) + 32f(x_3) + 7f(x_4)\right]$, $\text{Error} = \frac{8f^{(6)}(c)}{945}h^7$, $\text{Composite error} = \frac{2(b-a)f^{(6)}(c)}{945}h^6$

For the corresponding **composite** version of each rule, the weights (coefficients) of function values at some of the internal nodes are **doubled** due to simple summation of basic rules given above, as explained on preceding pages.

In addition, the **midpoint method** is an example of a so-called **open** Newton-Cotes method that does not include endpoints of the integration interval, but uses instead equidistant midpoints of each sub-partition:

$$\text{Midpoint method: } M_1 = h f(x_{mid}), \quad \text{Error} = -\frac{f''(c)}{24}h^3, \quad \text{Composite error} = -\frac{(b-a)f''(c)}{24}h^2$$

We derived the error of the midpoint method in class, using the same technique of checking monomials, $\boldsymbol{f(x) = x^n}$.

In all error expressions above, ***c*** is some point lying within the integration domain.