CHEMICAL ENGINEERING 2E04

Chapter 4 – Differentiation and Integration

Module 4B: Integration

Dr. Jake Nease
Kieran McKenzie
Steven Karolat
Cynthia Pham

Chemical Engineering
McMaster University

Updated November 17, 2019

Contents

Supplementary Material

Suggested Readings

Gilat, V. Subramaniam: *Numerical Methods for Engineers and Scientists.* Wiley. Third Edition (2014)

- Newton-Cotes Formulas
  - Chapter 9 → *9.3, 9.4,*
- Richardson Extrapolation
  - Chapter 9 → *9.9*
- Romberg Integration
  - Chapter 9 → *9.10*
- Evaluation of Multiple Integrals
  - Chapter 9 → *9.6*

Online Material

Trapezoid Rule Derivation PDF *(Avenue to Learn)*

- Derives the Trapezoid Rule formula by approximating $f(x)$ with a first-order Lagrange polynomial, and finding the integral of that polynomial

MIT Single Variable Calculus > Explanation of Simpson's Rule *(Click to follow hyperlink)*

- Derives the Simpson's 1/3 formula using graphical and logical reasoning

Introduction and Perspective

Just like with differentiation, we can approximate an integral using discrete data points from a graph by using numerical methods! No equation required.

We'll also look at how to numerically integrate curves that ARE defined by a known equation. Why would we use numerical methods to do this instead of find the *exact* integral of a known function (either by hand or by using `MATLAB`)? See below for the answer:

$f(x) = \int_3^6 (5x^3 + 2x^2 + 1)dx$	"EASY to integrate! So quick!"
$f(x) = \int_1^{24} \left(\frac{\sqrt{x^2+1} - \sqrt{x^2+1}}{\sqrt{x^4-1}} - \frac{\sec^2(x)}{9+\tan^2(x)} + x\tan^{-1}(x^2) \right) dx$	"Please don't ask me to integrate." ☹
$f(x) = \int_{-3}^{2} e^{-x^2} dx \quad f(x) = \int_{-\pi}^{\pi} \frac{\sin(x)}{x} dx \ , \quad f(x) = \int_{-7}^{62} \frac{1}{(e^x - x)} dx$	"Wait, these are actually impossible to integrate using elementary functions – *Symbolab* can't even do them!"

See my point? Many equations can take a long time to integrate analytically, and some are even impossible to do so! Thankfully, we have numerical methods to save us time and grief.

Primary Learning Outcomes

- Discuss the *Newton-Cotes' formulas* for integration.
- Recognize the *error* associated with the various Newton-Cote's formulas.
- Use *Richardson extrapolation* and *Romberg integration* to achieve improved integral estimates.

Applications of Numerical Integration

- Calculating total plant emissions flow rates
- Determining total energy transferred from temperature readings
- Newton's law of cooling
- Process Control (integral term)

Newton-Cotes' Formulas

Now we'll lead into the Newton-Cotes' (NC) Formulas for numerical integration. Good news – all of these formulas can be intuitively *derived by using Lagrange polynomials!*

Concept Behind NC Formulas

The basic concept of NC formulas is this:

- Begin with the function, OR discrete data points from a black-box system.
- Fit some Lagrange polynomial to that function or data.
- Integrate the Lagrange polynomial over the desired interval, $[a, b]$.
  - Adjustments can be made to achieve more exact integral approximations.

Each unique NC formula is formed by using a different order of Lagrange polynomial to fit the function or data.

Before we get started on the NC formulas, we will set a common definition for *panels* for integration.

Defining our Panel of Integration

When we integrate, we will have a panel of integration. This panel represents one evaluation of a given NC formula. The panel size requirement can vary between different NC methods, depending on the type of Lagrange polynomial the rule is derived from and depending on whether we have data points of a function equation!

We will define one panel as size of h, where $h = \frac{b-a}{n}$, where n is the number of panels we want to use over our interval.

Recall the Rectangle Methods from earlier – each 'panel' represented each rectangle of integration.

Trapezoid Rule

The first NC formula that we'll introduce is the Trapezoid Rule, which is *very similar* to the Rectangle Methods. Instead of using rectangles to approximate the area, this rule uses trapezoids (*gasp*).

The Trapezoid Rule represents a function using a *first order* Lagrange polynomial – this results in a *line* being formed between *two points*, which can be looked at as a trapezoid. The area of the resulting trapezoid will be the integral approximation.

Example: Using the Trapezoid Rule on a Set of Data

Here is a simple example of the Trapezoid Rule in action over an interval, $[a, b]$ using one panel with size h as well as a demonstration of how that 'trapezoid' can be split into a triangle and rectangle. The area of the trapezoid $(A_{rectangle} + A_{triangle})$ will represent the approximate integral of this data set over the interval specified:

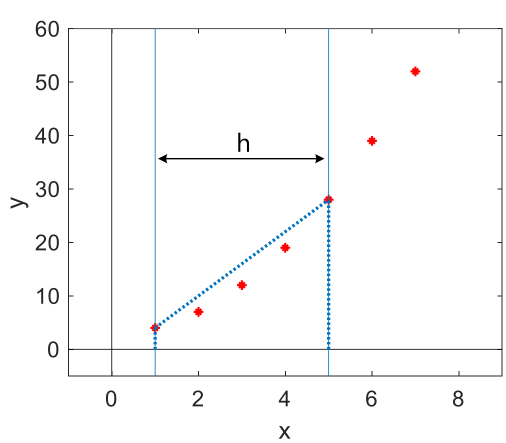


Figure 1: One panel of the Trapezoid Rule over the entire interval of integration $(1,5)$.

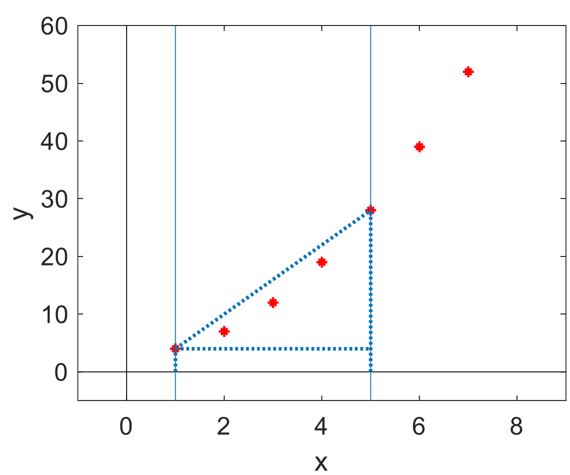


Figure 2: One panel of the Trapezoid Rule, shown split into a triangle and rectangle.

Approximated Area: 64

True Area: $53.\overline{3}$

Trapezoid Rule: Deriving the General Formula

Suppose we wish to find $\int_a^b f(x)dx$ numerically, using only data points or function evaluations rather than integrating the function analytically. To start, we'll assume that the interval $[a, b]$ only contains *one panel*. Then, representing $f(x)$ as a first order Lagrange polynomial:

$$\int_a^b f(x)dx \approx \int_a^b \left(\frac{(x-b)}{(a-b)} f(a) + \frac{(x-a)}{(b-a)} f(b) \right) dx$$

In this case, we can define our panel width as $h = (b-a)$ and $-h = (a-b)$. Substituting h into the equation above gives:

$$\int_a^b f(x)dx \approx \int_a^b \left(\frac{(b-x)}{h} f(a) + \frac{(x-a)}{h} f(b) \right) dx$$

After integrating and rearranging, you can turn this equation into a general formula below (the full derivation of the Trapezoid Rule has been included for you in the supplementary material. Take a peek!). The formula to numerically approximate an integral with the Trapezoid Rule becomes:

$$\boxed{\int_a^b f(x)dx \approx \frac{h}{2}\big(f(a) + f(b)\big)}$$

Trapezoid Rule (One Panel)

The following formula can be used to approximate the integral of a function over the interval $[a, b]$, where the panel size is defined as $h = b - a$ since we are only using 1 panel:

$$\int_a^b f(x)dx \approx \frac{h}{2}\big(f(a) + f(b)\big)$$

Notes:

- If used on discrete data points, the bounds a and b must be chosen at a location where the data points are known!
- For trapezoid rule, each panel fits a linear polynomial to two points.

We have the general formula for the Trapezoid Rule, using only two points. This creates *one trapezoidal panel* over an interval $[a, b]$, as shown in Figure 1. What if we wish to use multiple trapezoidal panels? For that, we must use the *composite Trapezoid Rule*.

Composite Trapezoid Rule

We can break the interval, $[a, b]$, into n panels over the integral limits $[a, b]$. In the example above, we have $N = 5$ *points* within $[a, b]$, and can therefore be split $[a, b]$ into $n = 4$ *panels!* Shown below is the Trapezoid Rule performed with 4 panels, again with an extra plot showing how those trapezoidal panels could be visualized as the superposition of triangles and rectangles:

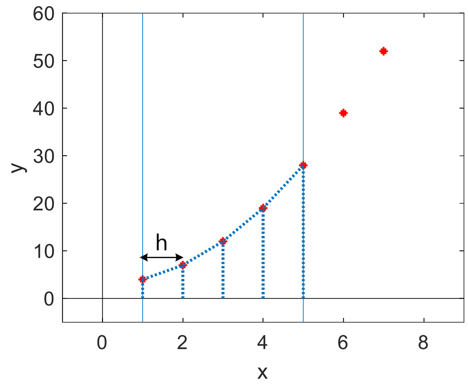


Figure 3: Trapezoid Rule shown using 4 *panels over the interval* $(1,5)$.

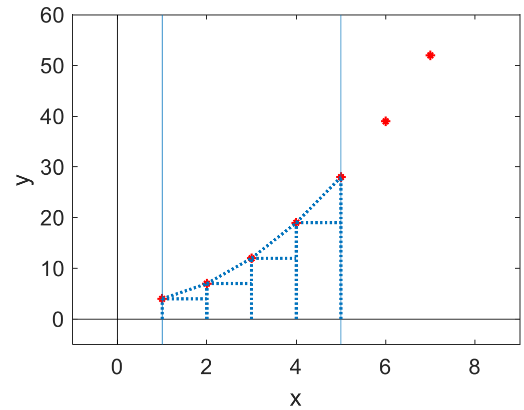


Figure 4: Trapezoid Rule shown using 4 *panels with the trapezoids split into a triangle and rectangle.*

Approximated Area: 54
True Area: $53.\overline{3}$

To use n panels, perform the general formula above over each panel to find the area of each trapezoid, and then sum them. In other words, we want:

$$\int_a^b f(x)dx \approx \underbrace{\frac{h}{2}(f(a) + f(a+h))}_{First\ Trapezoid} + \underbrace{\frac{h}{2}(f(a+h) + f(a+2h))}_{Second\ Trapezoid} + \ldots + \underbrace{\frac{h}{2}(f(a+(nh)) + f(b))}_{Final\ Trapezoid}$$

From here, there are two ways we could approach the composite trapezoid rule:

1. Common factor like terms to obtain a simplified formula:

$$\int_a^b f(x)dx \approx \frac{h}{2}\left(f(a) + 2f(a+h) + \cdots + 2f(a+(nh)) + f(b)\right)$$

Can be simplified to:

$$\boxed{\int_a^b f(x)dx \approx \frac{h}{2}\left(f(a) + 2\sum_{i=1}^{n}\{f(a+ih)\} + f(b)\right)}$$

Where $h = \frac{b-a}{n}$, and n is the number of panels used.

2. Leave the formula *as is* and perform single panel Trapezoid Rule n times (*i.e.* for n panels), summing the results as you go along.

NOTE: Our formula assumes h is constant for each use of the Trapezoid Rule, but this second option can be easily modified to accommodate non-constant step sizes.

Composite Trapezoid Rule

The following two approaches can be used to approximate the integral of a function over the integration limits $[a, b]$, where this interval is split into n panels chosen each with width $h = \frac{b-a}{n}$:

1. $$\int_a^b f(x)dx \approx \frac{h}{2}\left(f(a) + 2\sum_{i=1}^{n-1}\big(f(a+ih)\big) + f(b)\right)$$

2. $$\int_a^b f(x)dx \approx \sum_{i=1}^{n} \frac{h}{2}\big(f(a+(i-1)h) + f(a+ih)\big)$$

Notes:

- If used on discrete data points, the bounds a and b must be chosen at a location where the data points are known.
- These expressions assume constant spacing between data points.

Workshop: Coding the Trapezoid Rule

Complete the pseudocode below to code up the trapezoid method! Some of the lines have been filled in already. We'll throw it in MATLAB when we're done.

Pseudo Code	**MATLAB Code**
Calculate the number of panels (n).	`n = (b - a)/h;`
Create a dummy variable to start our integral.	`I = 0 ;`
We want to determine each of the n trapezoid areas	`for i = 1:n`
We need the right bound of the trapezoid	`b =`
Evaluate the function at left bound	`fa =`
Evaluate the function at right bound	`fb =`
Compute the area and add er' up	`I = I +`
Update the lower bound	`a =`
End the for-loop.	`end`

Question: The pseudo code above assumes that h is constant, and $[a, b]$ can be equally spaced by h. What if I have a set of data points which are UNEVENLY spaced, and want to use the Trapezoid Rule to approximate the integral?

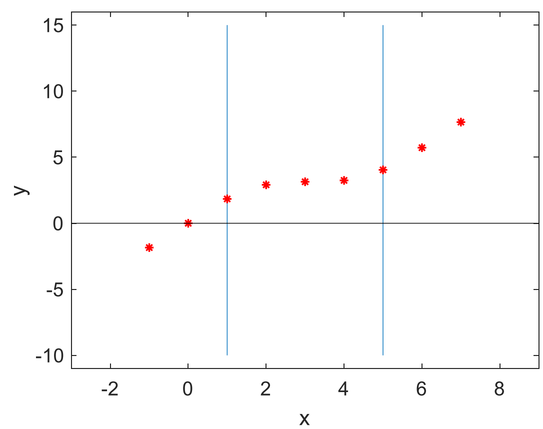
Answer: Look back at our derivation of the composite Trapezoid Rule formula, before 'like-terms' were grouped together. There you will find your answer!

Simpson's 1/3 Rule

This rule uses a *second order Lagrange Polynomial* to represent the function $f(x)$ on a given set of subintervals. The formula for this rule can be derived much like the Trapezoid Rule (skipped here), and will turn out to be:

$$\int_a^b f(x)dx \approx \frac{h}{6}\left(f(a) + 4f\left(a + \frac{h}{2}\right) + f(b) \right)$$

Where $h = \frac{b-a}{n}$, and in this case $n = 1$ for 1 panel. Note that Simpson's 1/3 Rule fits *quadratic polynomial* over the interval $[a, b]$, which requires the use of 3 points per panel evaluation. The data set below has 5 *points* over the interval $[a, b]$, and can therefore be fit to *up to* 2 *quadratic panels*. For now, we will show 1 panel, using 3 points, over the entire range for the Simpson's $1/3$ rule:



0

Figure 5: Example data set with 5 data points.

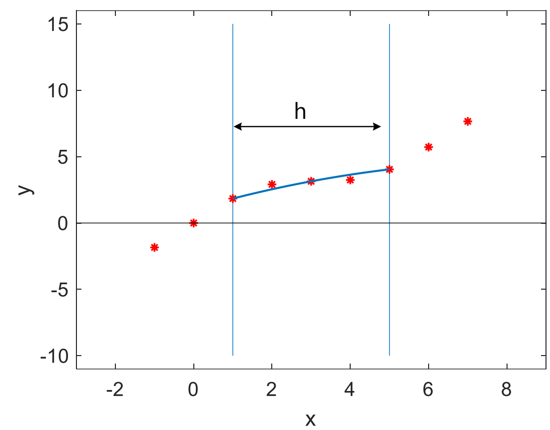


Figure 6: Simpson's $1/3$ rule applied to the example data using one panel.

Note: This data was formed by the function: $f(x) = sin(x) + x$

Approximated Area: 12.2980

True Area: 12.2566

When only fitting the data with one panel, note that *the median* point of our data within $[a, b]$ was used as the third data point (as you can see on the graph).

Simpson's 1/3 Rule (One Panel)

The following formula can be used to approximate the integral of a function, over integration limits $[a, b]$, where the step size is defined as $h = b - a$. Note that $n = 1$ for one panel:

$$\int_a^b f(x)dx \approx \frac{h}{6}\left(f(a) + 4f\left(a + \frac{h}{2}\right) + f(b)\right)$$

Notes:

- If used on discrete data points, the bounds a and b must be chosen at a location where the data points are known!
- Each panel fits a quadratic polynomial to three points.
- The distance between the data points used is required to be constant.

Composite Simpson's 1/3 Rule

In the exact same way as we derived the composite Trapezoid Rule, we can derive the formula for the Composite Simpson's 1/3 Rule, once again having two approaches. If the distance between data points is constant, and we use $h = \frac{b-a}{n}$ where n is the number of panels then we get:

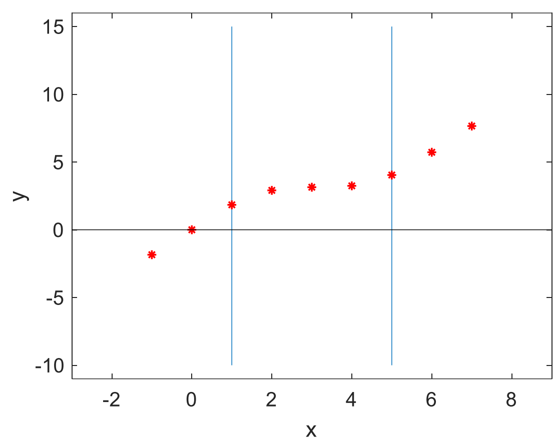


Figure 7: Example data with 5 data points.

Note: This data was formed by: $f(x) = sin(x) + x$.

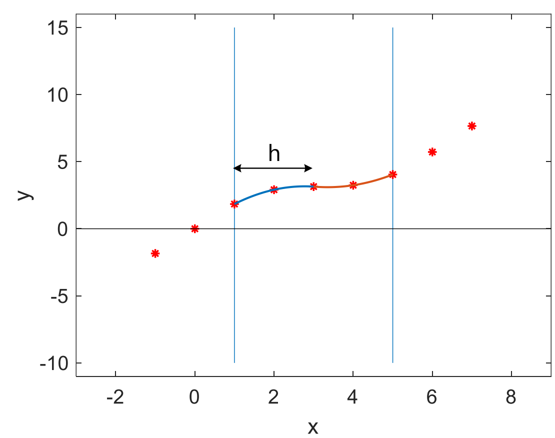


Figure 8: Composite Simpson's 1/3 Rule applied to the example data using two panels over the 5 data points.

Approximated Area: 12.2583

True Area: 12.2566

Composite Simpson's 1/3 Rule

The following methods can be used to approximate the integral of a function over the integration limits $[a, b]$, where $h = \frac{b-a}{n}$ and n is the number of panels:

1. $$\int_a^b f(x)dx \approx \frac{h}{6}\left(f(a) \ + \ 4\sum_{i=1}^{n}\left(f\left(a + (2i-1)\frac{h}{2}\right)\right) \ + \ 2\sum_{j=1}^{n-1}\left(f(a + (j)h)\right) \ + \ f(b)\right)$$

2. $$\int_a^b f(x)dx \approx \sum_{i=1}^{n} \frac{h}{6}\left(f(a + (i-1)h) + 4f\left(a + (2i-1)\frac{h}{2}\right) + f(a + (i)h)\right)$$

Notes:

- If used on discrete data points, the bounds a and b must be chosen at a location where the data points are known, and the spacing between data points must be constant.
- Each Simpson's 1/3 Rule panel uses *three* points.
  - Using the maximum number of panels improves the accuracy of the approximated integral.

Workshop: Pseudo Code for Simpson's 1/3 Rule

Draft out a function that will compute an approximate integral of a given function $f(x)$ using Simpson's 1/3 Rule on the interval $[a, b]$ using panel size h.

```
function I = simp13(f,lb,ub,h)









end
```

Simpson's 3/8 Rule

The last NC formula we'll talk about is the Simpson's 3/8 Rule, which can be derived using a *third order Lagrange Polynomial* to represent the function being integrated. The general formula for the approximated integral will be:

Simpson's 3/8 Rule (One Panel)

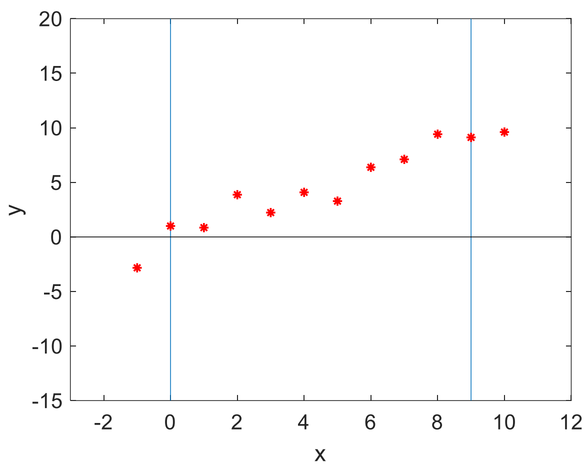
The following formula can be used to approximate the integral of a function between an interval a to b using four points, where once again, the panel size is defined as $h = b - a$ since we are using 1 panel:

$$\int_a^b f(x)dx \approx \frac{h}{8}\left(f(a) + 3f\left(a + \frac{h}{3}\right) + 3f\left(a + 2\frac{h}{3}\right) + f(b)\right)$$

Notes:

- If used on discrete data points, the bounds a and b must be chosen at a location where the data points are known!
- Each panel fits a cubic polynomial to 4 *points*.
- The distance between the data points used *is required* to be constant.

The data set at right has 10 *data points* over the interval $[a, b]$, and can therefore be split into 9 *subintervals*, for which, *up to* 3 *cubic panels* can be fit:



First, using the single Simpson's 3/8 Rule, we can fit a single curve to this data (Figure 9). Because we need 4 data points per panel, this interval can be fit perfectly with three panels instead of one, using the *composite Simpson's 3/8 Rule* (Figure 10).

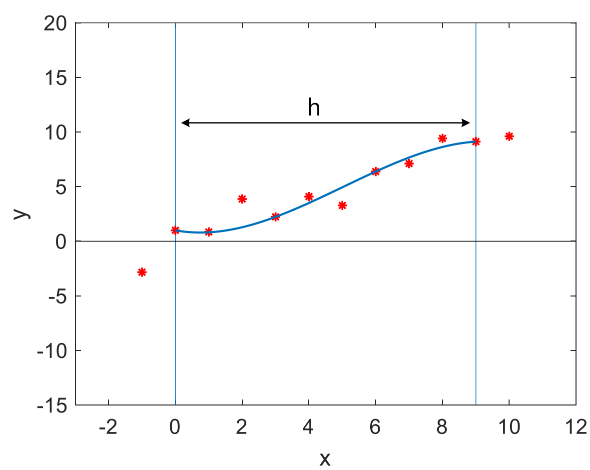


Figure 9: Simpson's $3/8$ *Rule for one panel with* 4 *points.*

Approximated Area: 40.4467

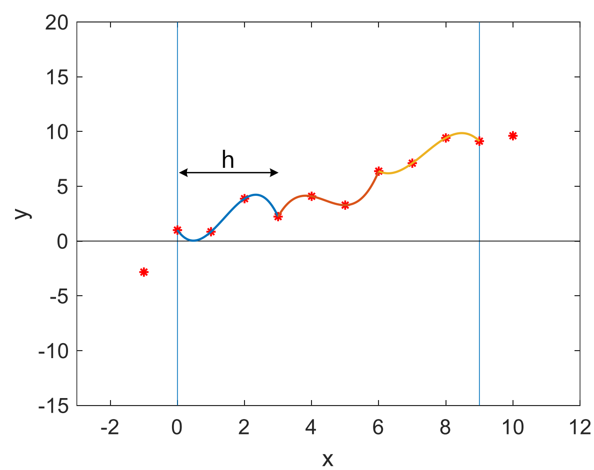


Figure 10: Composite Simpson's $3/8$ *rule using* 3 *panels with all points being used.*

Approximated Area: 42.4419

True Area: 42.7299

Composite Simpson’s 3/8 Rule

Composite Simpson’s 3/8 Rule

The following formulas can be used to approximate the integral of a function over the interval $[a, b]$, where this interval is split into n panels and $h = \frac{b-a}{n}$.

1. $$\int_a^b f(x)dx \approx \frac{h}{8}\left(f(a) + 3\sum_{i=1}^{n}\left(f\left(a + (3i-2)\frac{h}{3}\right) + f\left(a + (3i-1)\frac{h}{3}\right)\right) + 2\sum_{j=i}^{n-1}\left(f(a + (j)h)\right) + f(b)\right)$$

2. $$\int_a^b f(x)dx \approx \sum_{i=1}^{n}\frac{h}{8}\left(f(a + (i-1)h) + 3f\left(a + (3i-2)\frac{h}{3}\right) + 3f\left(a + (3i-1)\frac{h}{3}\right) + f(a + ih)\right)$$

Notes:

- If used on discrete data points, the bounds a and b must be chosen at a location where the data points are known, and the spacing between data points must be constant.
- Each Simpson’s 3/8 Rule panel uses four points.
  - Using the max number of panels improves the accuracy of the approximated integral.

NC Formulas: Errors

Like all numerical methods, NC formulas have error associated with them. Assuming we are using one panel for the entire range we get:

NC Formula	Panel Error (for some $\boldsymbol{\xi} \in [\boldsymbol{a}, \boldsymbol{b}]$)	Composite Error
Trapezoid Rule	$-\frac{h^3}{24}f^{(2)}(\xi) = \mathcal{O}(h^3)$	$\mathcal{O}(h^2)$
Simpson’s 1/3 Rule	$-\frac{h^5}{180}f^{(4)}(\xi) = \mathcal{O}(h^5)$	$\mathcal{O}(h^4)$
Simpson’s 3/8 Rule	$-\frac{3h^5}{80}f^{(4)}(\xi) = \mathcal{O}(h^5)$	$\mathcal{O}(h^4)$

From above, we can determine the *degree of precision (r)* for each method we have covered! For example, we can see that the error term for the Simpson’s 1/3 Rule depends on the $f^{(4)}(\xi)$, and will therefore give exact results (no error!) for all polynomials of $1^{st} - 3^{rd}$ degree ($\therefore r = 3$). Higher order NC formulas give a lower error term, however, risk overfitting the data – the Lagrange polynomial used must actually represent the trend of the data! They also require more points to be used in their formula, and therefore require more computation. If you have more data points, why not just fit more panels of a lower order rule?

Richardson Extrapolation

Richardson Extrapolation (RE) is a technique which takes in two estimates of an integral to achieve a third estimate which is more exact. If the two input estimates were of error $\mathcal{O}(h^2)$, then the result would be an estimate which has error $\mathcal{O}(h^4)$.

The general formula for RE uses two known estimates of an integral with the same error of $\mathcal{O}(h^s)$ to calculate a new estimate that has error either $\mathcal{O}(h^{s+1})$ or $\mathcal{O}(h^{s+2})$:

$$\boxed{I = \frac{2^s I_{2n} - I_n}{2^s - 1}}$$

Where I_n is the first approximated integral, calculated using n panels over $[a, b]$, and I_{2n} is the second, which uses $2n$ panels. For example, say you use the composite Trapezoid Rule once using 3 panels, and again using 6 panels. You will now know the values of I_{2n} and I_n, as well as the value for s from the error table introduced previously – you can then solve for the improved estimate, I.

Workshop: Deriving Richardson Extrapolation for Trapezoid Rule

Consider an integral I_1 computed using the trapezoid rule with some panel width h_1. Then, consider the integral I_2 performed over the same range but using twice as many panels, and thus a panel width $h_2 = \frac{1}{2}h_1$. Derive the formula to achieve an improved estimate I and show that the error for I is $\mathcal{O}(h_1^4)$.

Romberg Integration

Suppose that you used two integral estimates, I_n and I_{2n} to get *an improved estimate*. What if you also get I_{4n}, and use that with I_{2n} to get *another improved estimate*. Then you use your two improved estimates TO GET AN EVEN BETTER estimate?!

This is called Romberg Integration, and can be seen below in Figure 11, using the composite Trapezoid Rule. Each column from the figure represents a higher *level* of Romberg Integration, indicated by the *column index* number. The *row index* indicates the *integral number* within the current level of Romberg Integration:

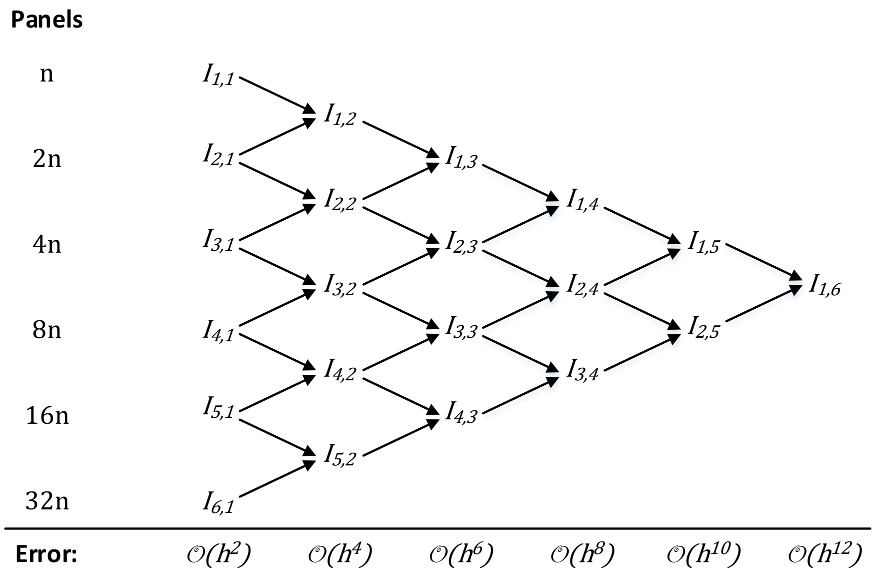


Figure 11: Romberg Integration using composite Trapezoid Rule. Powerful stuff!

Calculate the first column using the composite Trapezoid Rule. Then, to calculate a chosen extrapolated value from values in the previous level of Romberg Integration, the following formula is used:

$$\boxed{I_{i,j} = \frac{4^{j-1} I_{i+1,j-1} - I_{i,j-1}}{4^{j-1} - 1}}$$

Although we used the example of the composite Trapezoidal Rule, this process may be used with ANY composite rule covered.

Discussion: Adaptable Step Sizes

It is also possible to develop an *adaptable* step size algorithm that will check for consistency between step sizes and use a more refined step size if necessary. As a matter of fact, this can be used in a recursion loop. Let's discuss that in person and using `MATLAB`.

Conclusion and Summary

At last! Woohoo! In this module we have covered:

- A review of rectangle methods to find the integral of a function or data set.
- The various Newton-Cotes' formulas for solving an integral – Trapezoid Rule, Simpson's 1/3 Rule and Simpson's 3/8 Rule.
- Methods to improve our numerical integral estimation, such as Richardson extrapolation and Romberg integration.

To find out more, visit the *supplementary module* which covers:

- Integration using Gauss-Quadrature method
- Evaluation of multiple integrals (ex. Double integrals)

Next up: Differential Equations – The last chapter!

Appendix A: Trapezoid Rule Flow Chart