2020

Title: Numerical Integration with Different Methods

Question:

How does Simpson’s rule compare to the trapezium method when integrating power functions?

Table of Contents

1 Introduction 1
1.1 Trapezium Method 2
1.2 Simpson's Rule 3
1.3 Research Question 4
2 Investigation....................5
2.1 Error Depending on Coefficient and Power of the Power Function 6
2.1.1 Error Depending on Coefficient of the Power Function 8
2.1.2 Error Depending on Power of the Power Function.................... 10
2.1.2.1 Underestimate for Power between 0 and 1 11
2.1.2.2 Simpson's Rule Gives Exactly Integrates of Power Functions up to Cubic Degrees.................... 11
2.2 Error Depending on the Bounds of Integration 14
2.3 Error Depending on the Number of Segments 16
2.4 Modelling 19
3 Conclusion and Extension 25
Bibliography 27

1 Introduction

Integration is the inverse operation of differentiation. Where differentiation is used to find the change of a variable with respect to another at a specific point of a function by calculating the change in one variable when an infinitely small change in the other variable occurs. Integration is used to reverse this process to obtain the original function. Integrals can also provide expressions for the area under the curve of a function between two specified points in the case of definite integrals. Integration, has applications in many fields, such as engineering or computer science, in which it may be used to calculate volumes for example.

When finding integrals, approaches fall in two main categories: Analytical and numerical approaches. In analytical approaches, theorems are applied to find exact expressions for an integral, a solution is then calculated using the expression. In numerical approaches, algorithms are used to obtain a number solution[1]. Analytical methods will yield exact answers, whereas numerical methods will not always provide exact answers[2]. As such, analytical methods are not without their merit. There are, however, situations in which analytical methods are not applicable such as functions which cannot be integrated analytically e.g. $\sin(x^2)$. There are also situations in which analytical solutions could be too long or too difficult to obtain[3]. In these cases, we consider numerical methods. This study investigates two popular approximate methods for the integration, Simpson's rule and the trapezium method. These two methods are

---

[1] Jason Brownlee (2018). Analytical vs Numerical Solutions, <https://machinelearningmastery.com/analytical-vs-numerical-solutions-in-machine-learning/#:~:text=An%20analytical%20solution%20involves%20framing,solved%20well%20enough%20to%20stop.> [accessed 7 June 2020].
[2] Kendall E. Atkinson (2007). Numerical Analysis, <http://www.scholarpedia.org/article/Numerical_analysis> [accessed 7 June 2020].
[3] Maxwell Rosenlicht (1972). Integration in Finite Terms, American Mathematical Monthly. 79, pp. 963-972.

both ways to approximate definite integrals by estimating the area under the curve of a function and have been selected due to their simplicity and frequent use in teaching.

1.1 Trapezium Method

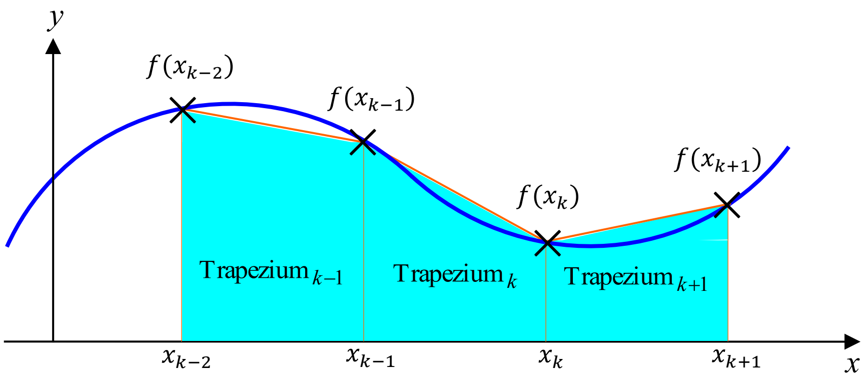

Figure 1 – A visualisation of the trapezium method

Given a function, the upper and lower bounds of the desired definite integral, and a limit to the number of trapeziums that may be drawn, the trapezium method uses horizontally equidistant points on the function, labelled in Figure 1. This is done depending on the limit to the number of trapeziums or segments, 3 in Figure 1. At each of these points, using the y-value trapeziums are drawn and their area is calculated.

The resulting mathematical expression in the context of Figure 1 is as follows:

$$I_{TR} = \sum_{k=1}^{n} \frac{f(x_{k-1}) + f(x_k)}{2} \Delta x_k \tag{1}$$

where $f(x_{k-1})$ is the y-value at point x_{k-1}, x_k is the subsequent x-value and Δx_k is the segment width.

1.2 Simpson's Rule

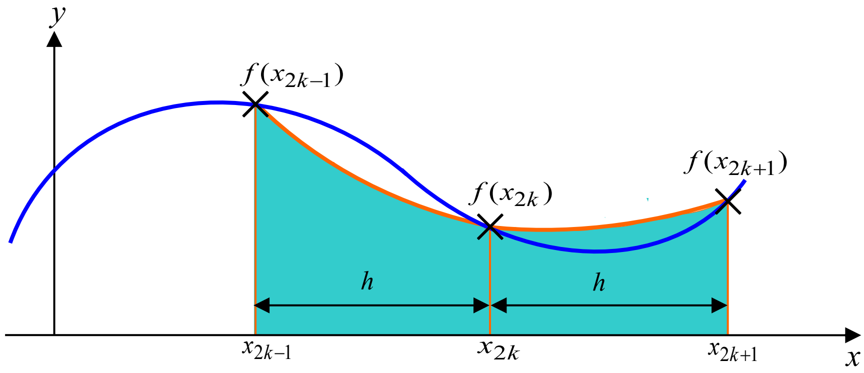

Figure 2 – A visualisation of Simpson's rule

Simpson's rule however, instead of using trapeziums tries to approximate a function with quadratic curves which can be analytically integrated[4]. Furthermore, to draw each of these curves, 3 points are used instead of 2. Using the 3 points, a start and end point are set and the middle point is used to set the turning point of the curve. Unlike the trapezium method which may work for any number of points, Simpson's rule only works for an odd number of points.

This series can be derived to produce the following expression:

$$I_{SR} = \frac{h}{3}\left[f(x_0) + 2\sum_{k=1}^{\frac{n}{2}-1} f(x_{2k}) + 4\sum_{k=1}^{\frac{n}{2}} f(x_{2k-1}) + f(x_n)\right] \tag{2}$$

where h is equal to the segment width Δx_k, n is the number of segments, each quadratic curve covers two segments, x_{2k} is the x-value at even intervals, and x_{2k-1} is the x-value at odd intervals.

[4] Milton Abramowitz and Irene A. Stegun (1972). Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables. 9th printing. New York: Dover, pp. 886.

1.3 Research Question

As stated before, the trapezium method divides the area under a curve into several trapeziums to calculate the area under the curve, using line segments to approximate a function of a curve and Simpson's rule uses quadratic curves. As a result, one might expect that the trapezium method will give a good approximation for functions which are closer to straight lines, or lines which have periods of mostly constant derivatives. Simpson's rule could be expected to better approximate functions with non-constant derivatives, e.g. variables of higher absolute values of powers[5].

Upon researching literature on Simpson's rule and the trapezium method, we can find that Simpson's rule and the trapezium method have been studied before[6]. Much academic writing already exists detailing Simpson's rule and the trapezium method[7, 8]. Writing mainly made comparisons between the methods with qualitative analyses without example cases. Furthermore, methods of error mitigation to improve these basic numerical methods have not been explored. Many queries still exist. For what kind of functions, if any will the trapezium method surpass Simpson's rule? How do the bounds of integration and number of segments used affect error? How do the coefficient and the power index of power function integrand affect error for example? Can the error of these methods be accurately predicted with a model? Being able to model the error of a function is valuable and can allow for mitigation of the error by addition or subtraction. In the field of computer science where a fluid might be being modelled using small square faces, i.e. trapezoidal prisms, being able to estimate error

---

[5] Harold Jeffreys and Bertha Swirles Jeffreys (1988). Methods of Mathematical Physics. 3rd ed. Cambridge, England: Cambridge University Press, pp. 286.
[6] Donald Kreider and Dwight Lahr (2002). Trapezoid Rule and Simpson's Rule, <https://math.dartmouth.edu/~m3cod/klbookLectures/406unit/trap.pdf> [accessed 29 May 2020].
[7] Kendall E. Atkinson (1989). An Introduction to Numerical Analysis. 2nd ed. John Wiley & Sons.
[8] Richard L. Burden and J. Douglas Faires (2000). Numerical Analysis. 7th ed. Brooks/Cole.

can provide feedback to a program to increase the number of prisms. To encompass these subsidiary questions, we can use the more general question, how does Simpson's rule compare to the trapezium method when integrating power functions? This study explores these factors and tries to model the error in relation to these for each method.

2 Investigation

To effectively devise experiments, variables which may affect the error of the two methods should be considered so that they may be tested for. We may first test the speculation made in Section 1.3 by integrating functions of constant and variable second derivatives with the two methods.

Power functions will be used. The relationships between different functions are clear, allowing for continuous data to be produced. Power functions are simple to integrate analytically and error may be calculated with ease. The functions are also limited to only one term in the form ax^b where a and b are real constants. The following integral will be investigated:

$$I = \int_{x_0}^{x_n} ax^b dx \tag{3}$$

This prevents overlaying trends from other higher and lower power terms which might respond differently to segment number change for example.

The number of segments, should also be controlled, ensuring data for each method is comparable. To reduce computational costs, a fairly low number of segments will also be used so that numerical integration errors can be relatively large and measurable. Ten segments will be used initially.

In the experiments, percentage error will be used to always keep error within context of the actual answers. In addition, we will not start incrementing power index b and coefficient a from exactly 0. Error at 0 will always be 0. Most importantly however, percentage error of the estimate cannot be calculated when the correct value of the definite integral is 0. These traits cause 0 to break what would be continuous data.

The increments of change to the power index will be below 1, such that any trends between integer powers can be seen and to allow better continuity in the data produced for a more conclusive analysis, i.e. no missed trends. The value chosen will be 0.2. This raises another issue, complex numbers will be produced if the power has an even denominator in its simplest form and one or more of the bounds of the integral are negative. Complex numbers do not find much relevance in the context of integration. As such negative bounds will be avoided.

Independent variables of the experiment need to be determined. From the power rule of differentiation and the fact that we are using power functions, it is clear there are two independent variables; the coefficient and the power of the power function. In addition, the error of the two methods will depend on the number of segments and may rely on the upper and lower bounds of integration as well. This essay will investigate the following relationships:

- Error depending on coefficient and power of the power function integrand
- Error depending on the bounds of integration
- Error depending on the number of segments

2.1 Error Depending on Coefficient and Power of the Power Function

In this experiment, the parameters used are listed in Table 1.

Table 1. The parameters used

Coefficient (a)	Lower bound (x_0)	Upper bound (x_n)	Power (b)	Number of segments (n)
1, 2, …, 50	0	1	0.2, 0.4, …, 10.2	10

As before, the function integrated is in the form ax^b. Eqs. (1), (2) and (3) can be written as follows:

$$I_{TR} = \sum_{k=1}^{10} \frac{ax_{k-1}^b + ax_k^b}{2} \Delta x_k$$

$$I_{SR} = \frac{h}{3}[2\sum_{k=1}^{4} ax_{2k}^b + 4\sum_{k=1}^{5} ax_{2k-1}^b + ax_{10}^b]$$

$$I = \int_0^1 ax^b dx$$

The trapezium method error can be calculated as:

$$E_{TR} = \frac{I_{TR} - I}{I} \times 100\% = \frac{\sum_{k=1}^{10} \frac{ax_{k-1}^b + ax_k^b}{2} \Delta x_k - \int_0^1 ax^b dx}{\int_0^1 ax^b dx} \times 100\%$$

Simpson's rule error can be calculated as:

$$E_{SR} = \frac{I_{SR} - I}{I} \times 100\% = \frac{\frac{h}{3}\left[2\sum_{k=1}^{4} ax_{2k}^b + 4\sum_{k=1}^{5} ax_{2k-1}^b + ax_{10}^b\right] - \int_0^1 ax^b dx}{\int_0^1 ax^b dx} \times 100\%$$

where h and Δx_k, the segment width, are equal to 0.1. x_0 is equal to zero, we have $x_k = x_{k-1} + \Delta x_k$.

The following figures are produced[9]:

[9] Jake VanderPlas (2016). Three-Dimensional Plotting in Matplotlib in Python Data Science Handbook, <https://jakevdp.github.io/PythonDataScienceHandbook/04.12-three-dimensional-plotting.html> [accessed 2 May 2020].

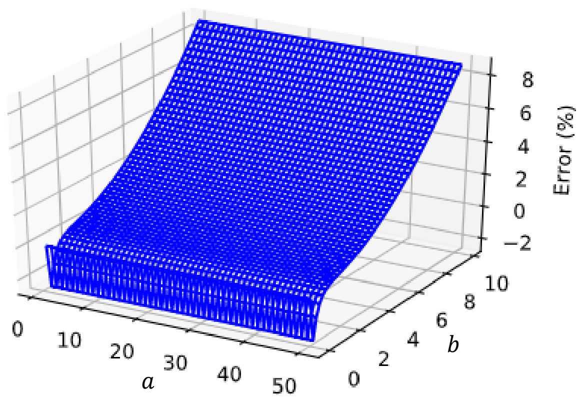

Figure 3 – Error of the trapezium method at different a, b values

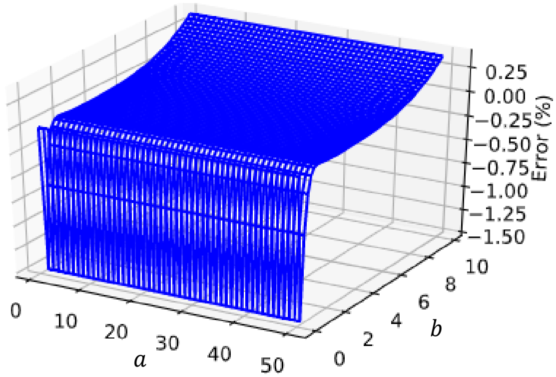

Figure 4 – Error of Simpson's rule at different a, b values

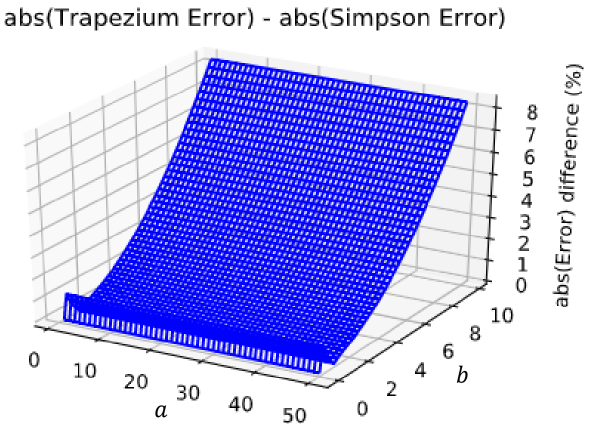

Figure 5 – Error difference of the trapezium method and Simpson's rule at different a, b values

The relationship between the error and power appears to be quite complex. Following an initial spike in error in which both methods appear to greatly underestimate an integral, both methods' inaccuracies increase at a growing rate as power increases. Simpson's rule has a much slower growing rate than the trapezium method. From Figure 5, contrary to earlier speculation, Simpson's rule surpasses or matches the trapezium method in accuracy for all powers.

2.1.1 Error Depending on Coefficient of the Power Function

As visible in Figures 3, 4 and 5, the coefficient actually has no effect on percentage error, a straight line is produced parallel to the coefficient axis. This does not change for different powers either. To explain this we can look to the percentage error expression.

The trapezium method error can be expressed as:

$$E_{TR} = \frac{\sum_{k=1}^{n} \frac{ax_{k-1}^{b} + ax_{k}^{b}}{2} \Delta x_k - \int_{x_0}^{x_n} ax^b dx}{\int_{x_0}^{x_n} ax^b dx} \times 100\%$$

Simpson's rule error can be expressed as:

$$E_{SR} = \frac{\frac{\Delta x_k}{3}\left[ax_0^b + 2\sum_{k=1}^{\frac{n}{2}-1} ax_{2k}^b + 4\sum_{k=1}^{\frac{n}{2}} ax_{2k-1}^b + ax_n^b\right] - \int_{x_0}^{x_n} ax^b dx}{\int_{x_0}^{x_n} ax^b dx} \times 100\%$$

The estimates and the actual integral are directly proportional to the coefficient, a. This means that the estimate error does increase with the coefficient. However, as percentage error is being used, the numerator and denominator in the above expressions are directly proportional to the coefficient, a, which can be cancelled from the above expressions to be rewritten as follows:

$$E_{TR} = \frac{\sum_{k=1}^{n} \frac{x_{k-1}^{b} + x_{k}^{b}}{2} \Delta x_k - \left[\frac{x^{b+1}}{b+1}\right]_{x_0}^{x_n}}{\left[\frac{x^{b+1}}{b+1}\right]_{x_0}^{x_n}} \times 100\% \qquad (4)$$

$$E_{SR} = \frac{\frac{\Delta x_k}{3}\left[x_0^b + 2\sum_{k=1}^{\frac{n}{2}-1} x_{2k}^b + 4\sum_{k=1}^{\frac{n}{2}} x_{2k-1}^b + x_n^b\right] - \left[\frac{x^{b+1}}{b+1}\right]_{x_0}^{x_n}}{\left[\frac{x^{b+1}}{b+1}\right]_{x_0}^{x_n}} \times 100\% \qquad (5)$$

As one can see, the coefficient a is no longer present in the percentage error expression which agrees with what was observed.

2.1.2 Error Depending on Power of the Power Function

While the above graphs provide an overview of how error changes relative to power and coefficient. To better view how error changes relative to power (b) the following figures are produced:

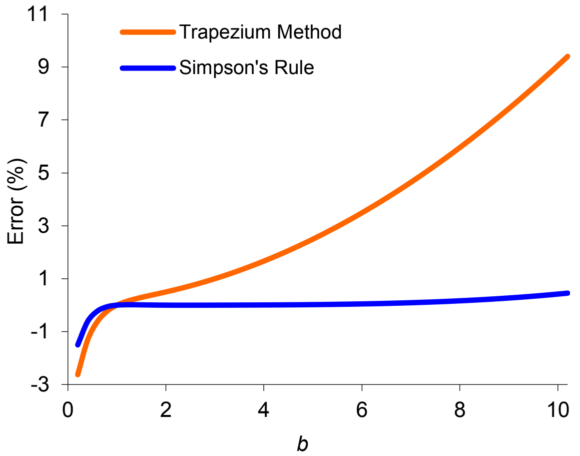

Figure 6 – Error against power index b

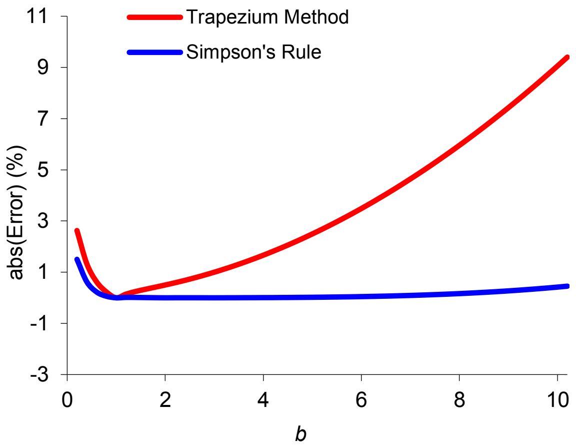

Figure 7 – Absolute error against power index b

Several trends may be observed; the initial spike in error which both methods appear to greatly underestimate an integral; the error decreases with the power index for both methods, when the power index, *b*, is between 0 and 1. This will be explained in Section 2.1.2.1.

When b is equal to 1, both methods produce no error. Most notably, Simpson's rule actually produces perfect definite integrals at powers 1, 2 and 3. Section 2.1.2.2 will explore this further.

The estimate error for the trapezium method starts to increase for b>1, while Simpson's rule starts to increase for b>3. Both methods appear to have a similar trend as power increases. This might be modelled by an equation, investigated in Section 2.4 below.

2.1.2.1 Underestimate for Power between 0 and 1

To understand the underestimate for both methods for power b between 0 and 1, $0 < b < 1$, the graph of the functions, $x^{0.1}$ and $x^{0.7}$, and trapeziums may be sketched as follows:

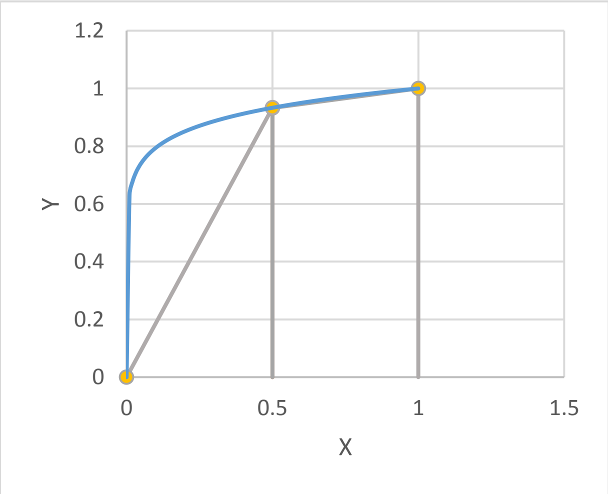

Figure 8 – Graph of function $y = x^{0.1}$ with trapeziums drawn

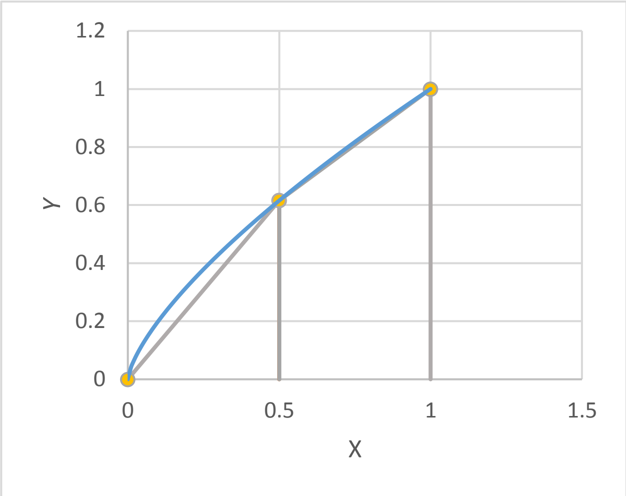

Figure 9 – Graph of function $y = x^{0.7}$ with trapeziums drawn

When b is very small, there is sudden growth and tapering at $x > 0$. The value of y at $x = 0$ is 0, however the value of y at $x > 0$ is close to 1. A large corner of the area is cut by the trapezium method, as shown in Figure 8. This corner cut results in initial spikes in error and an underestimate by using the trapezium method. As b increases, as shown in Figure 9, less of the area is cut off and the integral becomes more and more accurate until $b = 1$ is reached. Simpson's rule is similarly affected by the function shape.

2.1.2.2 Simpson's Rule Gives Exactly Integrates of Power Functions up to Cubic Degrees

Due to the trapezium method using line segments, it perfectly integrates linear functions. Simpson's rule produces perfect definite integrals at not only power 1, but 2 and 3 with errors of 0. Simpson's rule tries to approximate a function with quadratic curves. Based on the fact that given any three points, you can find the equation of a quadratic, or linear function when the coefficient of the squared term of the quadratic is 0, Simpson's rule can give perfect definite integrals with powers 1 and 2. It can be difficult to understand how Simpson's rule may also perfectly fit a cubic function. It can be mathematically proven as follows.

We define a cubic function as follows:

$f(x) = x^3$

n is the number of segments used to integrate, and h is the segment width.

$$\int_a^{a+nh} x^3\,dx = \left[\frac{x^4}{4}\right]_a^{a+nh} = \frac{(a+bh)^4 - a^4}{4}$$

$$= \frac{a^4 + 4a^3nh + 6a^2n^2h^2 + 4an^3h^3 + n^4h^4 - a^4}{4}$$

$$= \frac{nh(4a^3 + 6a^2nh + 4an^2h^2 + n^3h^3)}{4} \tag{6}$$

When Simpson's rule is used as follows:

$$\int_a^{a+nh} x^3\,dx \approx \frac{h}{3}\left[f(a) + 2\sum_{k=1}^{\frac{n}{2}-1} f(a+2kh) + 4\sum_{k=1}^{\frac{n}{2}} f(a+(2k-1)h) + f(a+nh)\right]$$

$$= \frac{h}{3}\left[a^3 + 2\sum_{k=1}^{\frac{n}{2}-1}(a+2kh)^3 + 4\sum_{k=1}^{\frac{n}{2}}(a+(2k-1)h)^3 + (a+nh)^3\right] \tag{7}$$

As

$$1^2 + 2^2 + 3^2 + \cdots + n^2 = \frac{n(n+1)(2n+1)}{6}, \qquad 1^3 + 2^3 + 3^3 \ldots + n^3 = \frac{n^2(n+1)^2}{4},$$

we have that

$$2\sum_{k=1}^{\frac{n}{2}-1}(a+2kh)^3 = 2\sum_{k=1}^{\frac{n}{2}-1}(a^3 + 6ka^2h + 12k^2ah^2 + 8k^3h^3)$$

$$= 2\sum_{k=1}^{\frac{n}{2}-1} a^3 + 2\sum_{k=1}^{\frac{n}{2}-1} 6ka^2h + 2\sum_{k=1}^{\frac{n}{2}-1} 12k^2ah^2 + 2\sum_{k=1}^{\frac{n}{2}-1} 8k^3h^3$$

$$= 2\left(\frac{n}{2} - 1\right)a^3 + \frac{3(n^2 - 2n)a^2h}{2} + (n^3 - 3n^2 + 2n)ah^2 + \frac{(n^4 - 4n^3 + 4n^2)h^3}{4} \quad (8)$$

As

$$1^2 + 3^2 + 5^2 \ldots + (2n-1)^2 = \frac{n(2n-1)(2n+1)}{3},$$

$$1^3 + 3^3 + 5^3 \ldots + (2n-1)^3 = n^2(2n^2 - 1),$$

we have that

$$4\sum_{k=1}^{\frac{n}{2}} (a + (2k-1)h)^3$$

$$= 4\sum_{k=1}^{\frac{n}{2}} (a^3 + 3(2k-1)a^2h + 3(2k-1)^2ah^2 + (2k-1)^3h^3)^3$$

$$= \sum_{k=1}^{\frac{n}{2}} 4a^3 + \sum_{k=1}^{\frac{n}{2}} 12(2k-1)a^2h + \sum_{k=1}^{\frac{n}{2}} 12(2k-1)^2ah^2$$

$$+ \sum_{k=1}^{\frac{n}{2}} 4(2k-1)^3h^3$$

$$= 4 \times \frac{n}{2} \times a^3 + \frac{\frac{12a^2h\left(1 + 2 \times \frac{n}{2} - 1\right)n}{2}}{2} + \frac{\frac{12ah^2(n+1)(n-1)n}{2}}{3}$$

$$+ 4h^3\left(\frac{n}{2}\right)^2\left(2 \times \left(\frac{n}{2}\right)^2 - 1\right)$$

$$= 2na^3 + 3n^2a^2h + 2(n^3 - n)ah^2 + \frac{(n^4 - 2n^2)h^3}{2} \quad (9)$$

$$(a + nh)^3 = a^3 + 3a^2nh + 3an^2h^2 + n^3h^3 \quad (10)$$

From Eqs. (8), (9) and (10), we have that

$$2\sum_{k=1}^{\frac{n}{2}-1} (a + 2kh)^3 + 4\sum_{k=1}^{\frac{n}{2}} (a + (2k-1)h)^3 + (a + nh)^3$$

$$= 3na^3 - a^3 + \frac{9}{2}n^2a^2h + 3n^3ah^2 + \frac{3n^4h^3}{4}$$

$$\frac{h}{3}\left[a^3 + 2\sum_{k=1}^{\frac{n}{2}-1}(a+2kh)^3 + 4\sum_{k=1}^{\frac{n}{2}}(a+(2k-1)h)^3 + (a+nh)^3\right]$$

$$= \frac{h}{3}\left[3na^3 + \frac{9}{2}n^2a^2h + 3n^3ah^2 + \frac{3n^4h^3}{4}\right]$$

$$= \frac{nh(4a^3 + 6a^2nh + 4an^2h^2 + n^3h^3)}{4}$$

This gives the same result as Eq. (6).

Therefore, Simpson's rule can give exact results for integrals of power functions up to cubic degrees.

2.2 Error Depending on the Bounds of Integration

Another variable which might affect the error of the definite integral would be the bounds of integral. Here the upper bound is increased from 1 to 50 to change the range of integration, lower bound is held constant at 0. The parameters used are listed in Table 2. The following figures are produced:

Table 2. The parameters used

Coefficient (a)	Lower bound (x_0)	Upper bound (x_n)	Power (b)	Number of segments (n)
1	0	1, 2, ..., 50	0.2, 0.4, …, 10.2	10

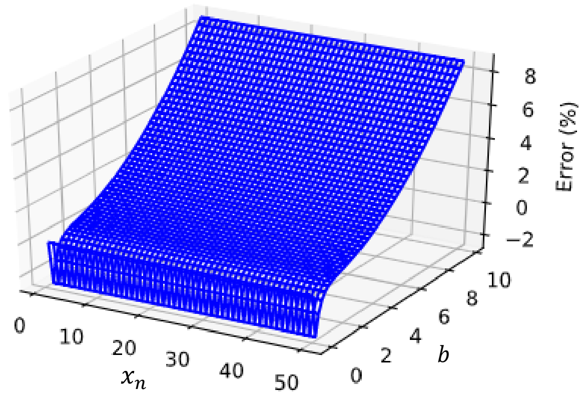

Figure 10 – Error of the trapezium method at different upper bounds and powers

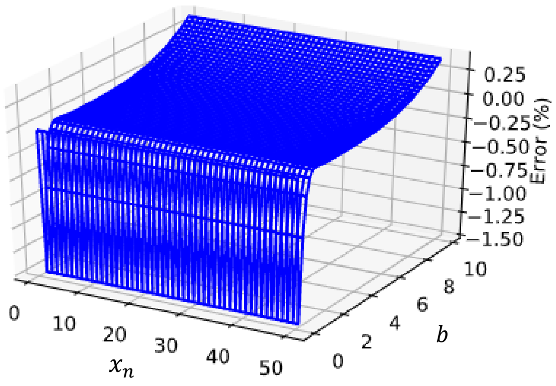

Figure 11 – Error of Simpson's rule at different upper bounds and powers

abs(Trapezium Error) - abs(Simpson Error)

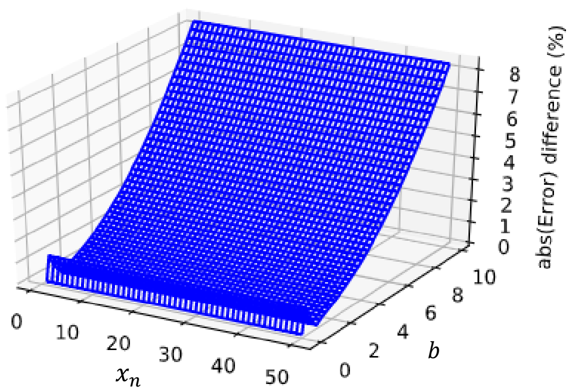

Figure 12 – Error difference of the trapezium method and Simpson's rule at different upper bounds and powers

As one would expect, error does increase as the upper bound increases however percentage error does not change. By increasing the bounds, we are increasing Δx_k, the segment width, leading to a larger error. Larger upper bounds however also increase the real integral value by a value proportional to the increase in error. There is no relationship between percentage error and the upper bound when lower bound is 0. It can be mathematically proven as follows:

For the given function $f(x) = ax^b$, the number of segments is m and the lower bound and upper bound are 0 and w, respectively, so that we have

$$\int_0^w ax^b dx = \left[\frac{ax^{b+1}}{b+1}\right]_0^w = \frac{aw^{b+1}}{b+1} \tag{11}$$

The trapezium method is used as follows.

$$\int_0^w ax^b dx \approx \frac{w}{2m}[2\sum_{k=1}^{m-1} f\left(\frac{kw}{m}\right) + f(w)] = \frac{w}{2m}[2\sum_{k=1}^{m-1} a(\frac{kw}{m})^b + aw^b] \tag{12}$$

When the upper bound increases from w to nw, Eqs. (11) and (12) can be written as follows:

$$\int_0^{nw} ax^b dx = \left[\frac{ax^{b+1}}{b+1}\right]_0^{nw} = \frac{a(nw)^{b+1}}{b+1} = \frac{aw^{b+1}n^{b+1}}{b+1} \tag{13}$$

$$\begin{aligned}\int_0^{nw} ax^b dx &\approx \frac{nw}{2m}\left[2\sum_{k=1}^{m-1} a(\frac{knw}{m})^b + a(nw)^b\right] \\ &= \frac{w}{2m}\left[2\sum_{k=1}^{m-1} a(\frac{kw}{m})^b + aw^b\right]n^{b+1}\end{aligned} \tag{14}$$

Comparing Eq. (13) to Eq. (11), and Eq. (14) to Eq. (12), the upper bound increases n times, the real integral value increases n^{b+1} times, and the estimate by the trapezium method also increases n^{b+1} times. Therefore the percentage error for the trapezium method is not changed by increasing the upper bound. Similarly, it can be mathematically proven that the percentage error for Simpson's rule is not changed by increasing the upper bound.

2.3 Error Depending on the Number of Segments

The following experiments are used to investigate how segment number affects the error for both methods.

Table 3. The parameters used

Coefficient (a)	Lower bound (x_0)	Upper bound (x_n)	Power (b)	Number of segments (n)
1	0	1	0.2, 0.4, …, 10.2	2, 4, 6, …, 40

The following three-dimensional figures only show the results for $b < 4$ for a better visualisation:

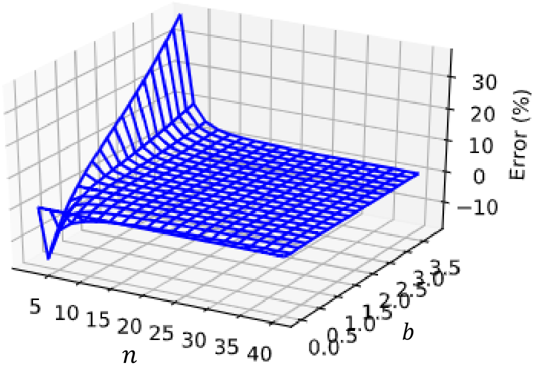

Figure 13 – Error of the trapezium method at different numbers of segments and powers

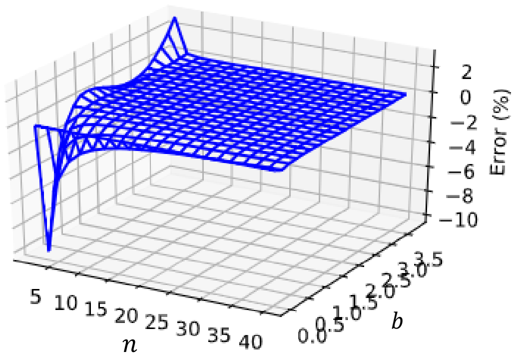

Figure 14 – Error of Simpson's rule at different numbers of segments and powers

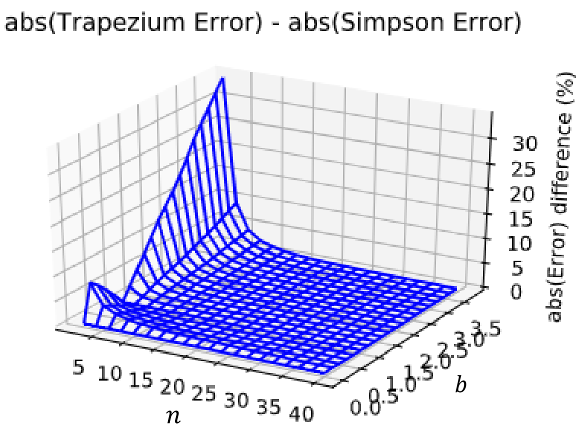

Figure 15 – Error difference of the trapezium method and Simpson's rule at different numbers of segments and powers

Figures 16 and 17 following display the detailed results for $b = 0.2, 2.2, 4.2, 6.2, 8.2$ and 10.2.

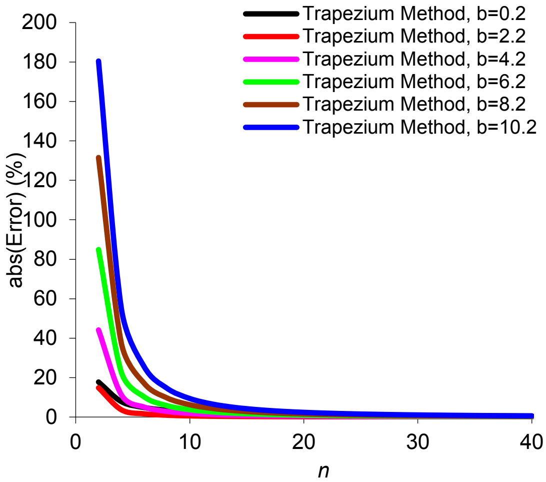

Figure 16 – Absolute error of the trapezium method at different numbers of segments for different powers

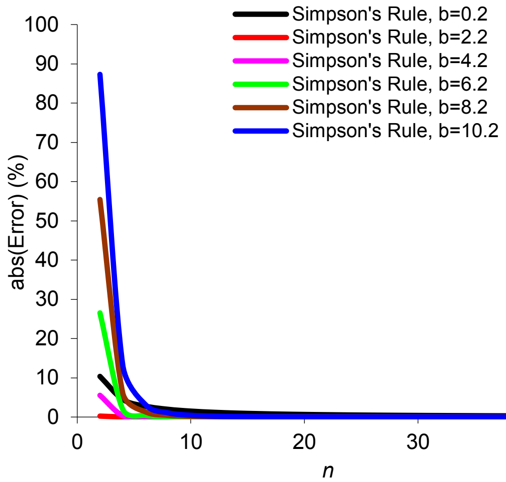

Figure 17 – Absolute error of Simpson's rule at different numbers of segments for different powers

To get a detailed view of Figures 16 and 17, they are redrawn as follows:

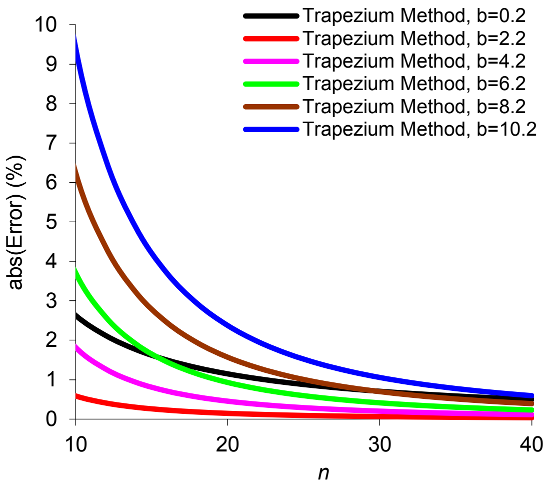

Figure 18 – Absolute error of the trapezium method at different numbers of segments for different powers

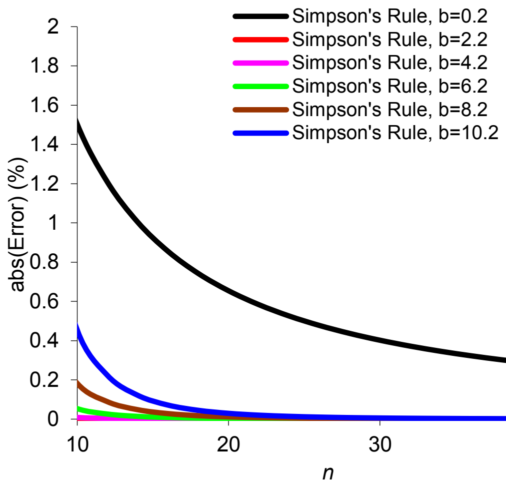

Figure 19 – Absolute error of Simpson's rule at different numbers of segments for different powers

As expected, both methods see dramatic declines in their error with segment increase. The error of both methods appear to be forming an asymptote where error is not equal to 0. As the number of segments tend to infinity, the error will tend to 0. This concept aligns with integration from first principles in which the area of some segments of a function are found and the number of segments n tend to infinity such that we achieve an exact answer. It should be noted that their estimate error decreases at a slowing rate.

If one were to calculate percentage change with segment number increase using the raw data, one would find Simpson's rule is in fact more strongly affected by an increase in the number of segments. This may be explained by how Simpson's rule is such that two segments are used at a time to approximate an integral, allowing smooth interpolation along three points depending on what fits best. Whereas the trapezium method which only takes into account one segment at a time interpolates linearly between two points. Simpson's rule makes better use of each segment in all situations. Simpson's rule can achieve the same results as the trapezium method with fewer segments.

2.4 Modelling

Section 2.3 shows that both methods see dramatic declines in their errors with segment increase from Figures 16, 17, 18 and 19. These declines actually follow power function patterns very closely.

Using power regression in Excel we can find the models which follow the relationship between the numerical method error (estimate error) and the number of segments, for a given power index, b. The expression for the trapezium method and Simpson's rule error models obtained would be as follows:

$E_{TR}(b,n) = cn^d$ (15) and $E_{SR}(b,n) = jn^m$ (16)

where n is the segment number and c, d, j and m are different values depending on b.

As c, d, j and m are dependent on power index, b, in the integrand, they are actually functions of b and can be modelled against b. Once the relationships are modelled, the error models, $E_{TR}(b, n) = cn^d$ and $E_{SR}(b, n) = jn^m$, can be obtained in terms of the power index, b and segment number, n.

To do this, values of c, d, j and m are found at particular b values by curve fitting the estimate error against the number of segments. This is repeated such that the dataset of c, d, j and m are obtained for different b values. From the dataset of c, d, j and m against b, models of c, d, j and m related to power index, b, can be found by using curve fitting. Error models can then be obtained using this information.

For example, when power index, b is equal to 10 for each method, the estimate error against the number of segments can be drawn, and then using curve fitting in Excel, the models for the error and number of segments relation are found to be:

$y = 763.57x^{-1.9428}$ for the trapezium method and $y = 1963.4x^{-3.7473}$ for Simpson's rule, as shown in Figure 20. From the fitted models, we can find the values of c, d, j and m at the given b, 763.57, -1.9428, 1963.4 and -3.7473, respectively, marked in red in Table 4. In the same way, the other values of c, d, j and m at the different values of b can be found and listed in Table 4.

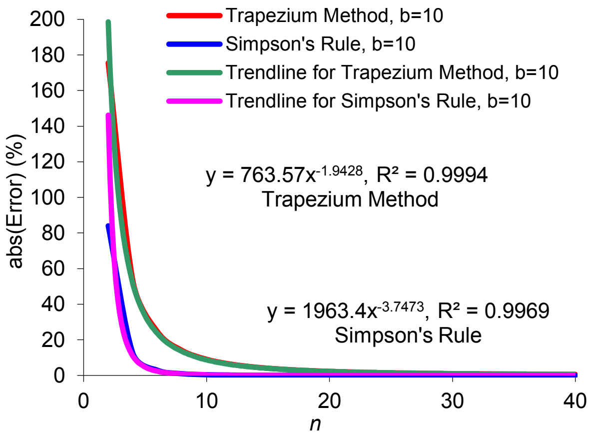

Figure 20 – Curve fitting the estimate error against number of segments at $b = 10$

Then, using the dataset, c, d, j and m can be expressed in terms of power index, b in models. The error models, $E_{TR}(b,n) = cn^d$ and $E_{SR}(b,n) = jn^m$, can then be obtained in terms of b and n as suggested earlier.

Table 4. Dataset of c, d, j and m obtained by the curve fitting for different power indexes, b

	Trapezium Method Error Model $E_{TR}(b,n) = cn^d$		Simpson's Rule Error Model $E_{SR}(b,n) = jn^m$	
b	c	d	j	m
1.2	11.088	-1.8928		
1.4	21.661	-1.9462	-	-
1.6	31.538	-1.9770	-	-
1.8	40.882	-1.9928	-	-
2	50.000	-2.0000	-	-
2.2	59.188	-2.0026	-	-
2.4	68.664	-2.0029	-	-
2.6	78.574	-2.0022	-	-
2.8	89.003	-2.0012	-	-
3	100.00	-2.0000	-	-
3.2	111.59	-1.9989	7.1461	-3.8317
3.4	123.77	-1.9978	17.674	-3.9023
3.6	136.56	-1.9967	31.329	-3.9510
3.8	149.93	-1.9957	47.746	-3.9821
4	163.87	-1.9947	66.667	-4.0000
4.2	178.37	-1.9937	88.015	-4.0087
4.4	193.41	-1.9926	111.86	-4.0113
4.6	208.97	-1.9916	138.36	-4.0098
4.8	225.04	-1.9904	167.69	-4.0058
5	241.59	1.9893	200.00	-4.0000
5.2	258.60	-1.9880	235.44	-3.9931
5.4	276.06	-1.9867	274.10	-3.9854
5.6	293.94	-1.9853	316.05	-3.9771
5.8	312.23	-1.9839	361.32	-3.9684
6	330.91	-1.9824	409.92	-3.9593
…	…	…	…	…
10	763.57	-1.9428	1963.4	-3.7473
10.2	787.27	-1.9406	2061.8	-3.7363

Using the above dataset in Table 4, the figures about c, d, j and m against the power index, b, in integrand, can be drawn. Figure 21 shows the relationship between d and the power index, b, for the trapezium method error. Figure 22 shows the relationship between m and the power index, b, for Simpson's rule error.

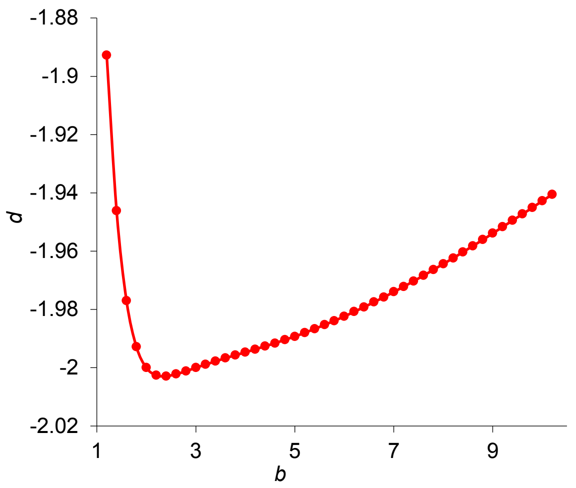

Figure 21 – d against power index *b* for the trapezium method

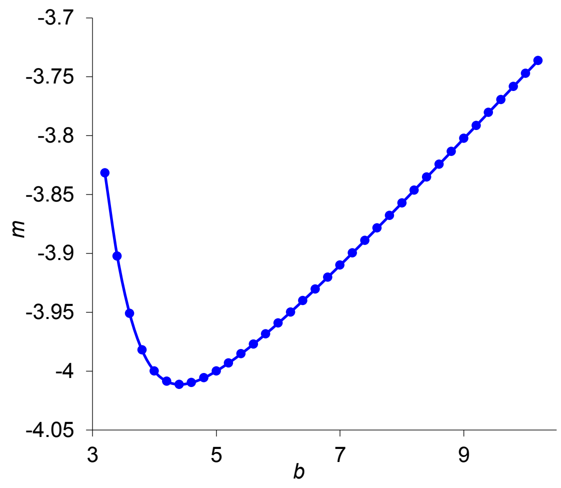

Figure 22 – m against power index *b* for Simpson's rule

From Figure 21, it can be seen that d against the power index, b, has two different patterns, one is that d decreases with the power index for $b < 2.4$, and the other is that d increases with the power index for $b > 2.4$. To simplify our modelling, we would have to separately model them.

Similarly, by looking at Figure 22, m against the power index, b, also has two different patterns, one is that m decreases with the power index for $b < 4.4$, and the other is that m increases with the power index for $b > 4.4$. Thus we would separately model these two patterns for Simpson's rule as well. Due to word constraints, this investigation will only model the error when the power index is greater than 2.4 for the trapezium method and when the power index is greater than 4.4 for Simpson's rule. Therefore, the figures about c, d, j and m against the power index, b, are drawn, as shown in Figures 23, 24, 25 and 26, respectively.

Figures 23, 24, 25 and 26 show that c, d, j and m, are increasing at an accelerating rate as power index increases. This indicates that the relationships between c, d, j, m and the power index, b are non-linear, suggesting that fitting polynomials with 2nd degrees will suffice as models.

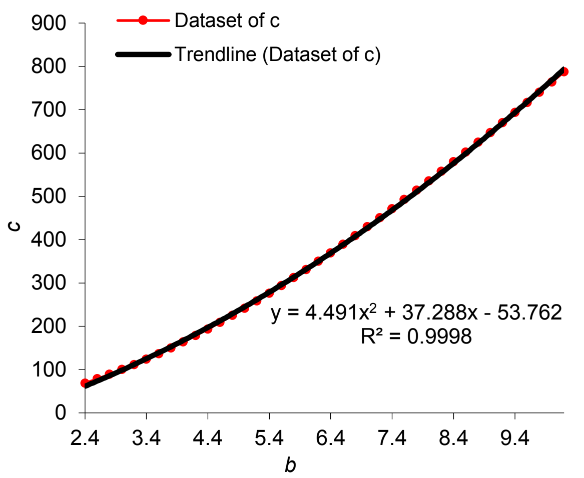

Figure 23 – c against power index b for the trapezium method

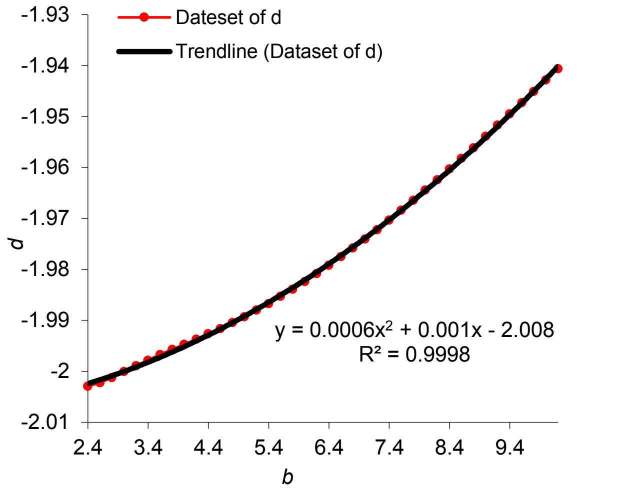

Figure 24 – d against power index b for the trapezium method

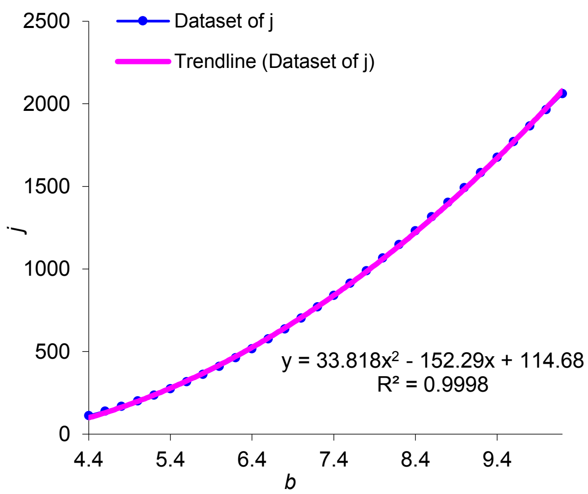

Figure 25 – j against power index b for Simpson's rule

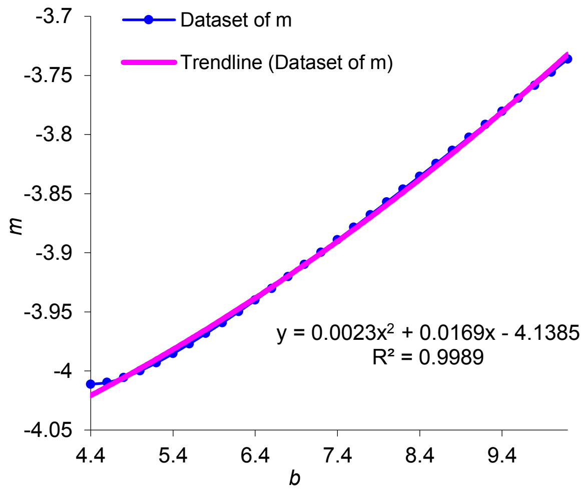

Figure 26 – m against power index b for Simpson's rule

From the best fitting model in Figure 23, c can be expressed as:

$$c = 4.491b^2 + 37.288b - 53.762$$

From the best fitting model in Figure 24, d can be expressed as:

$$d = 0.0006b^2 + 0.001b - 2.008$$

Therefore the error model for the trapezium method in Eq. (15) can be written as follows:

$$E_{TR}(b,n) = (4.491b^2 + 37.288b - 53.762)n^{0.0006b^2+0.001b-2.008} \quad (17)$$

Similarly, j from Figure 25 can be written as:

$j = 33.818b^2 - 152.29b + 114.68$

m from Figure 26 can be written as:

$m = 0.0023b^2 + 0.0169b - 4.1385$

Therefore the error model for Simpson's rule in Eq. (16) can be written as follows:

$$E_{SR}(b,n) = (33.818b^2 - 152.29b + 114.68)n^{0.0023b^2+0.0169b-4.1385} \quad (18)$$

Where b is the power index, and n is the number of segments.

Thus Eqs. (17) and (18) are the models for the error of the trapezium method and Simpson's rule, respectively, and have been produced.

Therefore the expression for the difference in error between the two methods is as follows:

$$E_{TR}(b,n) - E_{SR}(b,n) = (4.491b^2 + 37.288b - 53.762)n^{0.0006b^2+0.001b-2.008}$$
$$-(33.818b^2 - 152.29b + 114.68)n^{0.0023b^2+0.0169b-4.1385}$$

This provides a quantitative comparison of the two methods.

For a given power function integrand (b) and the number of segments (n), the error can be predicted by Eqs. (17) or (18). For a given power of the integrand (b) and the error (E), the number of segments (n) can be calculated by Eqs. (17) or (18), therefore the number of segments can be decided for a real experiment. Errors can also be estimated for functions which cannot be integrated analytically by approximating them as a series of power functions.

3 Conclusion and Extension

This investigation has shown with the use of mathematics, the relationships between variables such as bounds of an integral, the power index and coefficient of a power function, number of segments and the percentage error of two numerical methods of integration.

Some of the results and trends found were initially counterintuitive, however with further examination could be explained with the use of mathematics.

The upper bound has been demonstrated to affect the error in numerical integration using Simpson’s rule and the trapezium method linearly, if the lower bound is zero. Coefficient of the power function was shown to affect the error similarly.

When power index is greater than 0 and less than 1, the error decreases with the power index for both methods. When power index is equal to 1, they both produce zero errors. Most notably, Simpson’s rule actually produces perfect definite integrals at not only powers 0 and 1, but 2 and 3. Simpson’s rule surpasses or matches the trapezium method in accuracy for all powers.

Both methods see dramatic declines in their error when the segment number increases. Their error also increases greatly as power increases. Simpson’s rule makes better use of each segment in all situations, and can achieve the same results as the trapezium method with fewer segments.

This investigation successfully produced models for error depending on segment number given power function integrands. These findings showed clearly the relationship between Simpson's rule error and the trapezium method error for most power integrands and segment numbers. Findings may also be used to approximate the error of either method, allowing for better selection of segment number.

This investigation is limited in several ways. The investigation was limited to investigating power functions, which can only be relevant in specific circumstances. Furthermore, real world applications of the findings outside of knowledge were not explored in particular detail.

Possible areas of extension to this investigation could have been to investigate the error when integrating functions such as polynomials, exponential functions, trigonometric functions, etc. Alternative numerical models could have been trialled and compared such as Gaussian quadrature[10] and the Runge-Kutta method[11]. Given the time, these extensions would allow this investigation to find application for many more integrands.

[10] S. P. Venkateshan and Prasanna Swaminathan (2014). Numerical Integration Quadrat, <https://www.sciencedirect.com/topics/engineering/gaussian-quadrature-rule> [accessed 29 May 2020].
[11] Michael Zeltkevic (1998). Runge-Kutta Methods, <https://web.mit.edu/10.001/Web/Course_Notes/Differential_Equations_Notes/node5.html> [accessed 30 May 2020].

Bibliography

- Donald Kreider and Dwight Lahr (2002). Trapezoid Rule and Simpson's Rule, <https://math.dartmouth.edu/~m3cod/klbookLectures/406unit/trap.pdf> [accessed 29 May 2020].
- Harold Jeffreys and Bertha Swirles Jeffreys (1988). Methods of Mathematical Physics. 3rd ed. Cambridge, England: Cambridge University Press, pp. 286.
- Jake VanderPlas (2016). Three-Dimensional Plotting in Matplotlib in Python Data Science Handbook, <https://jakevdp.github.io/PythonDataScienceHandbook/04.12-three-dimensional-plotting.html> [accessed 2 May 2020].
- Jason Brownlee (2018). Analytical vs Numerical Solutions, <https://machinelearningmastery.com/analytical-vs-numerical-solutions-in-machine-learning/#:~:text=An%20analytical%20solution%20involves%20framing,solved%20well%20enough%20to%20stop.> [accessed 7 June 2020].
- Kendall E. Atkinson (2007). Numerical Analysis, <http://www.scholarpedia.org/article/Numerical_analysis> [accessed 7 June 2020].
- Kendall E. Atkinson (1989). An Introduction to Numerical Analysis. 2nd ed. John Wiley & Sons.
- Maxwell Rosenlicht (1972). Integration in Finite Terms, American Mathematical Monthly. 79, pp. 963-972.
- Michael Zeltkevic (1998). Runge-Kutta Methods, <https://web.mit.edu/10.001/Web/Course_Notes/Differential_Equations_Notes/node5.html> [accessed 30 May 2020].
- Milton Abramowitz and Irene A. Stegun (1972). Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables. 9th printing. New York: Dover, pp. 886.

- Richard L. Burden and J. Douglas Faires (2000). Numerical Analysis. 7th ed. Brooks/Cole.
- S. P. Venkateshan and Prasanna Swaminathan (2014). Numerical Integration Quadrat, <https://www.sciencedirect.com/topics/engineering/gaussian-quadrature-rule> [accessed 29 May 2020].