STEM Education

STEM Education, 3 (2): 130–147
DOI: 10.3934/steme.2023009
Received: 10 May 2023
Revised: 26 June 2023
Accepted: 30 June 2023
Published: 05 July 2023

https://www.aimspress.com/journal/steme

*Classroom note*

# Solving problems involving numerical integration (I): Incorporating different techniques

**William Guo***

School of Engineering and Technology, Central Queensland University, Bruce Highway, North Rockhampton, QLD 4702, Australia

* **Correspondence:** Email: w.guo@cqu.edu.au; Tel: +61-7-49309686.

Academic Editor: Dianshun Hu

**Abstract:** Numerical integration plays an important role in solving various engineering and scientific problems and is often learnt in applied calculus commonly through the trapezium and Simpson's methods (or rules). A common misconception for some students is that Simpson's method is the default choice for numerical integration due to its higher accuracy in approximation over the trapezium method by overlooking the requirement for using Simpson's method. As learning progressed to other numerical methods scheduled later in advanced mathematics, such as interpolations and computational modelling using computing tools like MATLAB, there is a lack of articulation among these numerical methods for students to solve problems solvable only by combining two or more approaches. This classroom note shares a few teaching and learning practices the author experienced in lectures, tutorials, and formal assessments on comparing or combining different numerical methods for numerical integration for engineering students in applied calculus and advanced mathematics over the past decade at Central Queensland University (CQU), a regional university in Australia. Each case represents a common concern raised or a mistake made by some students in different times. These efforts helped not only correct the misconception on the use of Simpson's method by some students, but also develop students' strategic thinking in problem solving, particularly involving decision-making for choosing the best possible method to produce a more appropriate solution to a problem that does not have an analytical solution.

**Keywords:** numerical methods, numerical integration, trapezium method, Simpson's method, error bound, interpolation, Maclaurin series, computing tools

## 1. Introduction

Numerical computing is useful for solving problems in engineering (and many other disciplines). Therefore, various numerical techniques are introduced to engineering students during their mathematics studies in applied calculus, advanced mathematics, and computational modelling in engineering curriculum in most universities in the world. For example, engineering students at Central Queensland University (CQU) of Australia used to learn solving nonlinear equations by Newton's method, numerical integration by the trapezium and Simpson's methods, and numerical approximation by Taylor's and McLaurin's series in applied calculus [1–3], curve fitting by interpolations and solving ODEs by Euler's and Runge-Kutta methods in advanced mathematics [4], followed by computational modelling using MATLAB [5].

In the normal sequence of mathematics teaching and learning, numerical integration is taught before interpolations and ODEs, so in most cases the trapezium and Simpson's methods are introduced to students at the time when learning applied calculus. Once progressed to the later stage to learn interpolations and modelling, numerical integration is no longer the focus of teaching and learning. As a result, the numerical techniques learnt separately earlier or later are often not incorporated to solve theoretical and applied problems to which analytical solutions are not available, for example $\int_0^1 \sqrt{1+x^4}\,dx$ . There is also a lack of seamless articulation between solving some integrals by analytical methods and by numerical methods for students in practice, which saw many students stuck with either the analytical methods or numerical methods for a problem solvable only by combining the two approaches. However, addressing these issues is not an easy task as it requires the instructor to recall the related topics previously learnt whenever progressing to a new topic later, and more importantly to demonstrate strategy and tactics of logical articulation with two or more techniques using meaningful examples to students.

This classroom note shares a few teaching and learning practices in comparing or combining different numerical methods for numerical integration the author experienced in lectures, tutorials, and formal assessments at CQU. Most cases presented in this note are reworked by the author from various examples delivered to engineering students in applied calculus and advanced mathematics involving numerical integration over the past decade. Each case represents a common concern raised or a mistake made by some students in different times.

The first three examples showed that the higher accuracy in approximation for numerical integration resulted from Simpson's method when the requirement is met may lead students to take Simpson's method as their default choice for approaching numerical integration, often overlooked the requirement that Simpson's method only applies to cases where the equal subdivision must have even numbers. Applying Simpson's method to problems with equal subdivision of odd numbers would produce a result much worse than that by the trapezium method. This misperception on using Simpson's method as the default choice occurred in a formal assessment showed in Example 4 where some students directly applied the method to nine data sets. In other special circumstances, the trapezium method can also overperform Simpson's method, which is demonstrated in Example 5.

The next three cases were triggered by students' uncertainty on how to assess accuracy of an estimated result from applying a numerical integration method where the analytical solution is not available, despite availability of error bound for the chosen method. These examples aimed to show the students: 1) how to choose the suitable method and interval size with respect to the given accuracy or error limit; 2) how to use an alternative method to verify the result produced by the chosen method; 3) how to incorporate numerical methods to solve scientific and engineering problems.

These examples helped develop students' strategic thinking in problem solving, particularly involving decision-making for choosing the best possible method to produce a more appropriate solution to a problem that does not have an analytical solution. Such ability is vital for engineering students because most real-world engineering problems may not have exact solutions; hence choosing an approximate result that is more appropriate to the circumstances is a decision engineers must make in real-world engineering projects.

In the rest of this paper, the trapezium and Simpson's methods are first outlined in Section 2 with five examples to demonstrate the strengths and weaknesses of the two most popular methods in numerical integration. Section 3 presents three more examples used in the past lectures and tutorials for engineering students to show how different techniques can be incorporated to solve problems involving integration. This classroom note is closed by discussions and conclusions presented in Section 4.

## 2. The trapezium and Simpson's methods for numerical integration

### 2.1. The trapezium method or trapezium rule

The trapezium method is to divide the given range into *n* vertical strips of equal interval and each strip is treated as a trapezoid (Figure 1). Hence, the subarea of a single strip should be

$$A_i = \frac{1}{2}(y_{i-1} + y_i)h \quad \text{where} \quad h = \frac{b-a}{n}.$$

Adding the *n* subareas together, the integral sought is expressed as

$$\int_a^b f(x)dx = \int_a^b ydx = \frac{h}{2}[y_0 + 2(y_1 + y_2 + y_3 + ... + y_{n-1}) + y_n]. \quad (1)$$

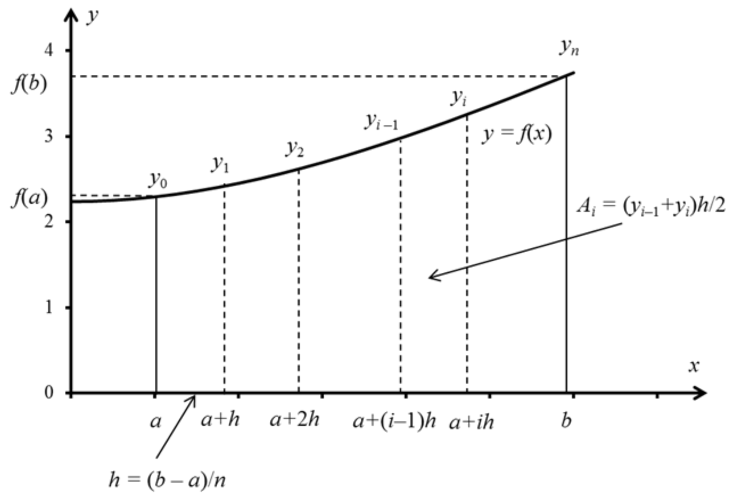


**Figure 1.** Sketch of the trapezium method.

Assuming the integrand $y = f(x)$ is continuous and its second-order derivative exists in range $[a, b]$, the error of a single strip and the total error of all strips together for approximation using the trapezium method were proven to be bounded by [6–9]

$$\begin{cases} E_i \le \dfrac{(b-a)^3 M}{12} \to O(h^3) \\ E_n \le \dfrac{(b-a)^3 M}{12n^2} \to O(h^2) \end{cases}, \tag{2}$$

where $M$ is the maximum value of the second-order derivative of $f(x)$ where $x = \xi$ is within the range, i.e., $|f''(\xi)| \le M,\ \xi \in [a,b]$. This indicates that the maximum error of the composite trapezium method is about the order of $O(h^2)$ even though that for a single strip is in the order of $O(h^3)$. If the width of a single interval is around 1/10 = 0.1, the maximum errors would be around the order of $O(10^{-3})$ for a single strip and of $O(10^{-2})$ for all $n$ strips together. Hence in general, the smaller the interval, the more accurate the approximation.

### 2.2. Simpson's method or Simpson's 1/3 rule

If dividing the range into vertical strips of equal interval by an *even-number*, a local quadratic interpolation can be created using the three known points of any two adjunct strips (Figure 2) to replace $y = f(x)$ within the subrange $[x_{i-1}, x_{i+1}]$. The area under the interpolation in $[x_{i-1}, x_{i+1}]$ is regarded as an approximate to the area under $f(x)$ for this subrange. Sliding this window through the whole range should get the areas for all paired vertical strips and their sum can be regarded as an approximate to the integral.

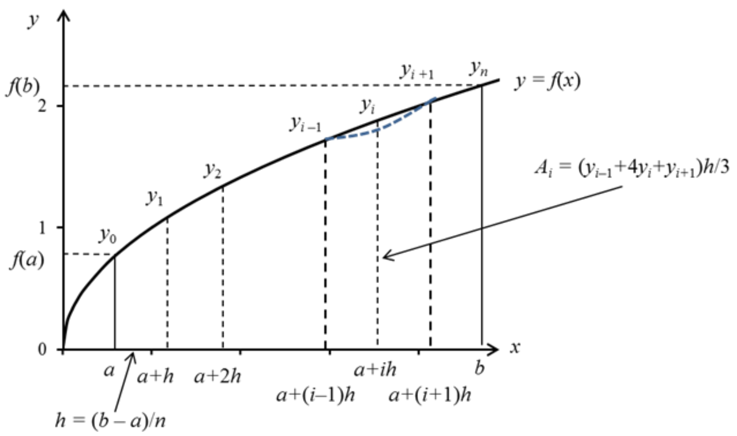


**Figure 2.** Sketch of the Simpson's method.

As the process is repeated by a sliding window of two adjacent strips over the whole range, i.e., from $x_0$–$x_2$ to $x_2$–$x_4$ to $x_4$–$x_6$ and so on, the formula of a subarea derived from any two adjacent strips should be applicable to other adjacent strips. In general, the subarea of any two adjacent strips centred at $x_i$ can be approximated by [10]:

$$S_i = \frac{h}{3}(y_{i-1} + 4y_i + y_{i+1}). \tag{3}$$

The total area of all strips can be obtained by adding all $n/2$ subareas together as follows:

$$\int_a^b ydx = \frac{h}{3}[y_0 + 4(y_1 + y_3 + ...) + 2(y_2 + y_4 + ...) + y_n]. \tag{4}$$

This is known as Simpson's method or Simpson's 1/3 rule for numerical integration [10–13].

Assuming the integrand $y = f(x)$ is continuous and its fourth-order derivative exists in range [$a$, $b$], the error of a single subarea and that of the total area approximated by Simpson's formulas (3-4) were proven to be bounded by [6,7,9]

$$\begin{cases} E_i \le \dfrac{(b-a)^5 M}{2880} \to O(h^5) \\ E_n \le \dfrac{(b-a)^5 M}{180n^4} \to O(h^4) \end{cases}, \tag{5}$$

where $M$ is the maximum value of the fourth-order derivative of $f(x)$ where $x = \xi$ is within the range, i.e., $\left|f^{(4)}(\xi)\right| \le M,\ \xi \in [a,b]$. Formula (5) indicates that the maximum error of the composite Simpson's method is about the order of $O(h^4)$. If the size of the interval is around 1/10 = 0.1, the maximum error would be in the order of $O(10^{-4})$ for all $n$ strips together, much more accurate than that of the trapezium method in general cases, which is demonstrated by the following examples. In all the examples presented in this paper, the numerical results are tabulated using Excel with a default truncation error to five decimal places to ensure that the final result is accurate to the fourth decimal place after rounding, unless the error limit is specified for a given problem.

If the range is divided into vertical strips of equal interval by a number that is a multiple of 3, a section of three consecutive strips can form a cubic polynomial that can be integrated analytically to obtain the area of this section. Moving this integral window across the whole range would form a composite formula expressed as follows, which is commonly called Simpson's 3/8 rule [12],

$$\int_a^b ydx = \frac{3h}{8}[y_0 + 3(y_1 + y_2 + y_4 + ...) + 2(y_3 + y_6 + ...) + y_n] = \frac{3h}{8}\left[y_0 + 3\sum_{i=1, i\ne 3k}^{n-1} y_i + 2\sum_{i=1}^{n/3-1} y_{3i} + y_n\right]. \tag{6}$$

Under the same condition, Simpson's 3/8 rule has an error bound in the order of $O(h^4)$ but about three times smaller than that of Simpson's method [12]. Other numerical methods for integration can be derived by using different orders of polynomials or combinations through the Newton-Cotes formulas [14]. However, the trapezium and Simpson's (or 1/3 rule) methods are the most used techniques in numerical integration.

**Example 1**: Use the trapezium method with seven strips of equal interval and Simpson's method with six strips of equal interval to approximate $\int_1^2 \frac{\ln x}{\sqrt{x}} dx$.

**Solution**

We use a table by Excel to realise the trapezium method (1) with $h$ = 1/7 as shown in the table below.

| $i$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Sum |
|---|---|---|---|---|---|---|---|---|---|
| $x_i$ | 1 | 1.14286 | 1.28571 | 1.42857 | 1.57143 | 1.71429 | 1.85714 | 2 | |
| $y_0$ or $y_n$ | 0 | | | | | | | 0.49013 | *0.49013* |
| $2y_i$ | | 0.24981 | 0.44328 | 0.59683 | 0.72112 | 0.82333 | 0.90850 | | *3.74287* |
| Integral | | | $h$×Sum/2 = **0.30236** | | | | | | ***4.23300*** |

Similarly, we can also use a table by Excel to realise Simpson's method (4) with six strips or $h$ = 1/6 as follows.

| $i$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | Sum |
|---|---|---|---|---|---|---|---|---|
| $x_i$ | 1 | 1.16667 | 1.33333 | 1.50000 | 1.66667 | 1.83333 | 2 | |
| $y_0$ or $y_n$ | 0 | | | | | | *0.49013* | *0.49013* |
| $4y_i$(odd) | | 0.57086 | | 1.32424 | | 1.79064 | | *3.68575* |
| $2y_i$(even) | | | 0.49828 | | 0.79137 | | | *1.28965* |
| Area | | | $h$×Sum/3 = **0.30364** | | | | | ***5.46553*** |

However, Simpson's method (4) cannot be applied to the case with seven strips where the trapezium method (1) was applied. Otherwise, the result would have a larger error that is demonstrated in the table blow.

| $i$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Sum |
|---|---|---|---|---|---|---|---|---|---|
| $x_i$ | 1 | 1.14286 | 1.28571 | 1.42857 | 1.57143 | 1.71429 | 1.85714 | 2 | |
| $y_0$ or $y_n$ | 0 | | | | | | | *0.49013* | *0.49013* |
| $4y_i$(odd) | | 0.49963 | | 1.19366 | | 1.64666 | | | *3.33995* |
| $2y_i$(even) | | | 0.44328 | | 0.72112 | | 0.90850 | | *2.07290* |
| Integral | | | $h$×Sum/3 = **0.28109** | | | | | | ***5.90298*** |

This integral has an exact solution [15]

$$\int_1^2 \frac{\ln x}{\sqrt{x}}\,dx = 2\sqrt{x}(\ln x - 2)\Big|_1^2 = 0.30366\,.$$

With this analytical solution as a reference, the absolute errors with respect to the exact solution from the trapezium and Simpson's methods with 7 and 6 subdivisions are tabulated below, along with the result from the misused Simpson's method to the seven strips. The error for the trapezium method is in the order of O($10^{-3}$) whereas the error of Simpson's method is in the order of O($10^{-5}$) even with 6 wider strips compared to the 7 narrower strips for the trapezium method. If the accuracy of approximation is set to be accurate to the fourth decimal place, Simpson's method would be the only choice for this case.

| Method | Approximate solution | Absolute error | Actual error $O(10^{-n})$ |
|---|---|---|---|
| Trapezium ($n$ = 7) | 0.30236 | 0.00130 | $10^{-3}$ |
| Simpson ($n$ = 6) | ***0.30364*** | 0.00002 | $10^{-5}$ |
| Misused Simpson ($n$ = 7) | 0.28109 | 0.02257 | $10^{-2}$ |
| Exact solution | | **0.30366** | |

An interesting observation on this example is the effect of the odd (7) and even (6) numbers for equal divisions over range [1, 2] on the results. In theory, the trapezium method applies to equal divisions of any number whereas Simpson's method only applies to equal divisions of even numbers. Due to higher accuracy of Simpson's method over the trapezium method in 'common' cases where the integrand is continuous and smooth, users may take Simpson's method as the default choice for numerical integration by neglecting the requirement on equal divisions of even numbers for applying Simpson's method. Should Simpson's method be applied to the case with seven equal intervals, the result would be with an error in the order of $O(10^{-2})$, worse than the order of $O(10^{-3})$ from the trapezium method.

The integrand of this question is continuous and its second and fourth derivatives exist as

$$y'' = \frac{3\ln x - 8}{4x^{5/2}} \text{ and } y^{(4)} = \frac{105\ln x - 352}{16x^{9/2}}.$$

The absolute maximum value for the second derivative in [1, 2] should be at $x$ = 1, i.e., $M = |y''(1)| = 2$. By formula (2), the error bound for the trapezium method with seven intervals would be

$$E_n \le \frac{(b-a)^3 M}{12n^2} = \frac{2}{12\times 7^2} \approx 0.003401 = 3.401\times 10^{-3}.$$

The actual error of 0.00130 by the trapezium method is indeed smaller than this estimated error bound.

The absolute maximum value for the fourth derivative in [1, 2] should be at $x$ = 1, i.e., $M = |y^{(4)}(1)| = 22$. By formula (5), the error bound for Simpson's method with six intervals would be

$$E_n \le \frac{(b-a)^5 M}{180n^4} = \frac{22}{180\times 6^4} \approx 9.431\times 10^{-5}.$$

The actual error of 0.00002 by Simpson's method is also smaller than this estimated error bound.

**Example 2**: Use the trapezium method with five strips of equal interval and Simpson's method with four strips of equal interval to approximate $\int_1^2 x^3 \ln x dx$.

**Solution**

We use a table by Excel to realise the trapezium method (1) with $h$ = 1/5 = 0.2 as a table below.

| $i$ | 0 | 1 | 2 | 3 | 4 | 5 | Sum |
|---|---|---|---|---|---|---|---|
| $x_i$ | 1 | 1.2 | 1.4 | 1.6 | 1.8 | 2 | |
| $y_0$ or $y_n$ | 0 | | | | | 5.54518 | *5.54518* |
| $2y_i$ | | 0.6301033 | 1.84656 | 3.85027 | 6.85594 | | *13.18288* |
| Integral | | | $h$×Sum/2 = **1.87281** | | | | ***18.72805*** |

Similarly, we can also use a table by Excel to realise Simpson's method (4) with four strips or $h$ = ¼ = 0.25 as follows.

| $i$ | 0 | 1 | 2 | 3 | 4 | Sum |
|---|---|---|---|---|---|---|
| $x_i$ | 1 | 1.25 | 1.5 | 1.75 | 2 | |
| $y_0$ or $y_n$ | 0 | | | | 5.54518 | *5.54518* |
| $4y_i$(odd) | | 1.74331 | | 11.99676 | | *13.74007* |
| $2y_i$(even) | | | 2.73689 | | | *2.73689* |
| Integral | | | $h$×Sum/3 = **1.83518** | | | ***22.02214*** |

However, should Simpson's method (4) be applied to the case with five strips, the result would have a larger error that is demonstrated in the table blow.

| $i$ | 0 | 1 | 2 | 3 | 4 | 5 | Sum |
|---|---|---|---|---|---|---|---|
| $x_i$ | 1 | 1.2 | 1.4 | 1.6 | 1.8 | 2 | |
| $y_0$ or $y_n$ | 0 | | | | | 5.54518 | 5.54518 |
| $4y_i$(odd) | | 1.26021 | | 7.70054 | | | *8.96075* |
| $2y_i$(even) | | | 1.84656 | | 6.85594 | | *8.70250* |
| Integral | | | $h$×Sum/3 = **1.54723** | | | | ***23.20843*** |

This integral has an exact solution [15]

$$\int_1^2 x^3 \ln x dx = \frac{1}{16} x^4 (4 \ln x - 1) \Big|_1^2 = 1.83509 .$$

With this analytical solution as a reference, the absolute errors with respect to the exact solution from the trapezium and Simpson's methods with 5 and 4 subdivisions are tabulated below, along with the result from the misused Simpson's method to the five strips. The error for the trapezium method is about the order of O($10^{-2}$) whereas that of Simpson's method is in the order of O($10^{-4}$) to O($10^{-5}$) even with 4 wider strips compared to the 5 narrower strips for the trapezium method. Should Simpson's procedure be applied to the five strips, the result would be with an error in the order of $O(10^{-1})$, even worse than the trapezium method. If the error limit of approximation is set to be better than parts per thousand (‰), Simpson's method would be the only choice for this case.

| Method | Approximate solution | Absolute error | Actual error $O(10^{-n})$ |
|---|---|---|---|
| Trapezium ($n$ = 5) | 1.87281 | 0.03772 | $10^{-2}$ |
| Simpson ($n$ = 4) | ***1.83518*** | 0.00009 | $10^{-5}$ |
| Misused Simpson ($n$ = 5) | 1.54723 | 0.28786 | $10^{-1}$ |
| Exact solution | | **1.83509** | |

The integrand of this question is continuous and its second and fourth derivatives exist as

$$y'' = 6x \ln x + 5x \text{ and } y^{(4)} = \frac{6}{x} .$$

The absolute maximum value for the second derivative in [1, 2] should be at $x$ = 2, i.e., $M = |y''(2)| = 18.31777$. By formula (2), the error bound for the trapezium method with five intervals would be

$$E_n \le \frac{(b-a)^3 M}{12n^2} = \frac{18.31777}{12 \times 5^2} \approx 0.06106 = 6.106 \times 10^{-2} .$$

The actual error of 0.03772 by the trapezium method is smaller than this estimated error bound.

The absolute maximum value for the fourth derivative in [1, 2] should be at $x$ = 1, i.e., $M = |y^{(4)}(1)| = 6$. By formula (5), the error bound for Simpson's method with four intervals would be

$$E_n \le \frac{(b-a)^5 M}{180n^4} = \frac{6}{180 \times 4^4} \approx 0.00013 = 1.3 \times 10^{-4} .$$

The actual error of 0.00009 by Simpson's method is also smaller than this estimated error bound.

**Example 3**: Use the trapezium method with seven strips of equal interval and Simpson's method with six strips of equal interval to approximate $\int_0^1 x \arctan x dx$.

**Solution**

We use a table by Excel to realise the trapezium method (1) with $h$ = 1/7 as below.

| $i$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Sum |
|---|---|---|---|---|---|---|---|---|---|
| $x_i$ | 0 | 0.14286 | 0.28571 | 0.42857 | 0.57143 | 0.71429 | 0.85714 | 1 | |
| $y_0$ or $y_n$ | 0 | | | | | | | 0.78540 | *0.78540* |
| $2y_i$ | | 0.04054 | 0.15903 | 0.34705 | 0.59331 | 0.88607 | 1.21479 | | *3.24079* |
| Integral | | | | $h$×Sum/2 = **0.28758** | | | | | ***4.02619*** |

The result using Simpson's method (4) with $h$ = 1/6 is shown in the following table.

| $i$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | Sum |
|---|---|---|---|---|---|---|---|---|
| $x_i$ | 0 | 0.16667 | 0.33333 | 0.50000 | 0.66667 | 0.83333 | 1 | |
| $y_0$ or $y_n$ | 0 | | | | | | 0.78540 | *0.78540* |
| $4y_i$(odd) | | 0.11010 | | 0.92730 | | 2.31579 | | *3.35319* |
| $2y_i$(even) | | | 0.21450 | | 0.78400 | | | *0.99850* |
| Integral | | | | $h$×Sum/3 = **0.28539** | | | | ***5.13709*** |

If Simpson's method were misused to the seven strips, the result would be as follows.

| $i$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Sum |
|---|---|---|---|---|---|---|---|---|---|
| $x_i$ | 0 | 0.14286 | 0.28571 | 0.42857 | 0.57143 | 0.71429 | 0.85714 | 1 | |
| $y_0$ or $y_n$ | 0 | | | | | | | 0.78540 | *0.78540* |
| $4y_i$(odd) | | 0.08108 | | 0.69410 | | 1.77214 | | | *2.54733* |
| $2y_i$(even) | | | 0.15903 | | 0.59331 | | 1.21479 | | *1.96713* |
| Integral | | | | | $h$×Sum/3 = **0.25237** | | | | ***5.29985*** |

This integral has an exact solution [16]

$$I = \int_0^1 x \arctan x dx = \frac{1}{2}[(x^2+1)\arctan x - x]\Big|_0^1 = 0.28540 .$$

With this analytical solution as a reference, the absolute errors with respect to the exact solution from the trapezium and Simpson's methods with 7 and 6 subdivisions are tabulated below, along with the result from the misused Simpson's method to the seven strips. Again, Simpson's method produced the best result with an error in the order of $O(10^{-5})$ for the six strips, much better than the trapezium method for the seven strips with an error in the order of $O(10^{-3})$. However, if Simpson's method were applied to the seven strips, the result would be the worst with an error in the order of $O(10^{-2})$.

| Method | Approximate solution | Absolute error | Actual error $O(10^{-n})$ |
|---|---|---|---|
| Trapezium ($n$=7) | 0.28758 | 0.00219 | $10^{-3}$ |
| Simpson ($n$=6) | ***0.28539*** | 0.00001 | $10^{-5}$ |
| Misused Simpson ($n$=7) | 0.25237 | 0.03302 | $10^{-2}$ |
| Exact solution | | **0.28540** | |

The integrand of this question is continuous and its second and fourth derivatives exist as

$$y'' = \frac{2}{(1+x^2)^2} \quad \text{and} \quad y^{(4)} = \frac{8(5x^2-1)}{(1+x^2)^4} .$$

The absolute maximum value for the second derivative in [0, 1] should be at $x$ = 0, i.e., $M = |y''(0)| = 2$. By formula (2), the error bound for the trapezium method with seven intervals would be

$$E_n \le \frac{(b-a)^3 M}{12n^2} = \frac{2}{12\times 7^2} \approx 0.003401 = 3.401\times 10^{-3} .$$

The actual error of 0.00219 by the trapezium method is indeed smaller than this estimated error bound.

The absolute maximum value for the fourth derivative in [0, 1] should be at $x$ = 0, i.e., $M = |y^{(4)}(0)| = 8$. By formula (5), the error bound for Simpson's method with six intervals would be

$$E_n \le \frac{(b-a)^5 M}{180n^4} = \frac{8}{180\times 6^4} \approx 3.429\times 10^{-5}.$$

The actual error of 0.00001 by Simpson's method is also smaller than this estimated error bound.

All the examples above demonstrated that under similar (not the same) conditions, Simpson's method would produce a more accurate approximate than the trapezium method. This led some users to use Simpson's method as their default choice to approach problems involving numerical integration by neglecting the basic requirement that Simpson's method is only applicable to the cases with sequential datasets of equal interval with even numbers. The following example was from an assignment to the past engineering students.

**Example 4**: A plot of land lies between a straight fence ($x$-axis) and a stream (northern bound). Measured from the western end of the fence, the breadth of the plot ($y$-axis) was recorded in the table below. Choose an appropriate method to estimate the area of this plot of land. Keep 1 decimal place in the final result.

| $x$ (metre) | 0 | 3 | 6 | 9 | 12 | 15 | 18 | 21 | 24 | 27 |
|---|---|---|---|---|---|---|---|---|---|---|
| $y$ (metre) | 16.3 | 17.9 | 20.7 | 22.8 | 23.7 | 23.3 | 21.9 | 19.8 | 18.5 | 19.7 |

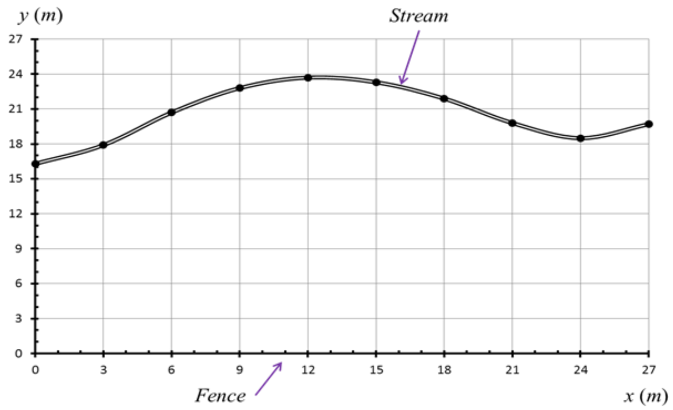


**Figure 3**. Plot of the measurements for the block of land in Example 4.

**Solution**

This question was assigned to 20 engineering students as part of a formal assessment several years ago. Since there are nine equal intervals over 27 metres, the trapezium method would be a more appropriate choice than Simpson's method. The estimated area for this problem should be 559.8 $m^2$ by directly using the trapezium method.

All the 20 students obtained the same or a similar value by applying the trapezium method. However, eleven students also directly used Simpson's method to approximate the area with a different value of 540.8 $m^2$ as shown in Figure 4, for which they could not explain why such a large difference in the land area was resulted from the two methods.

Three students did not use Simpson's method as they correctly noted that Simpson's method was unsuitable to this problem with 9 datasets. Other six students correctly identified the unsuitability of directly using Simpson's method to this problem but also proposed to combine the trapezium and Simpson's methods together for this case. They applied Simpson's method to the first 8 strips and the trapezium rule to the 9th strip and finally added the two together as an estimate to the land area (Figure 5). The area of 559.9 $m^2$ by this modified method is almost the same as that from the trapezium method.

**Simpson's Rule**

$$A = \frac{h}{3}(y_0 + y_n + 4\sum(y_{\text{odd}}) + 2\sum(y_{\text{even}}))$$
$$= \frac{3}{3}[(16.3) + (19.7) + 4(20.7) + 4(23.7) + 4(21.9) + 4(18.5) + 2(17.9) + 2(22.8) + 2(23.3) + 2(19.8)]$$
$$= 540.8\text{m}^2$$

**Figure 4**. An example of misusing Simpson's method to estimate the land area in Example 4.

**Using the Simpsons Rule:**

$$where\ n \neq odd\ number\ and\ must\ be\ an\ even\ number,$$

As n = odd number the Simpsons rule cannot be used to calculate the entire area under the curve. Instead by using the Simpsons Rule for $0 \geq i \geq 8$ (even number) and adding the trapezium rule for $8 \geq i \geq 9$ the area can be accurately found.

$$A_{under\ curve} \approx \frac{h}{3}(y_0 + y_{n-1} + 4\sum('odd'heights) + 2\sum('even'heights)) + \frac{h}{2}(y_n - y_{n-1})$$

$$\approx \frac{3}{3}(16.3 + 18.5 + 4(17.9 + 22.8 + 23.3 + 19.8) + 2(20.7 + 23.7 + 21.9)) + \frac{3}{2}(19.7 + 18.5)$$

$$\approx (502.6) + \frac{3}{2}(38.2)$$

$$\approx 502.6 + 57.3$$

$$\approx 559.9m^2$$

**Figure 5**. An example of using the modified method to estimate the land area in Example 4.

Since this problem has nine known datasets, a multiple of three, Simpson's 3/8 rule (6) can be used to check the estimated results from the two methods. The result from Simpson's 3/8 rule would be 559.2 $m^2$, which is very close to the estimated area by both the trapezium and modified methods used by the students.

To answer a recent query from a student about whether the trapezium method would perform better than Simpson's method under the same condition, the following example was conceptualised based on the recent numerical modelling with a self-balanced two-wheel robot [17].

**Example 5**: A self-balanced two-wheel robot was controlled by combination of a constant acceleration and a 'periodic turbulence' during the first 10 seconds. The two components for the speed of the motion were defined as follows:

$$\begin{cases} v_1 = 0.5t & 0 \leq t \leq 10s \\ v_2 = |t-1| & p = 2s, t \in [0,10] \end{cases}.$$

Choose an appropriate method to estimate the accumulated distance the robot travelled in the first 10 seconds. Keep 1 decimal place in the final result.

**Solution**

The speed at any time during this period should be the sum of $v_1$ and $v_2$. Using an equal interval of 1 second, ten datasets can be calculated and are tabulated below. A corresponding *t*-*v* diagram is drawn in Figure 6.

| *t* (s) | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| *v* (m/s) | 1.0 | 0.5 | 2.0 | 1.5 | 3.0 | 2.5 | 4.0 | 3.5 | 5.0 | 4.5 | 6.0 |

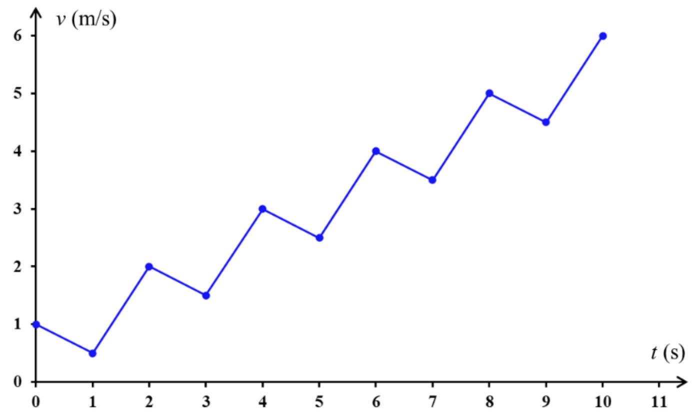


**Figure 6**. The *t*-*v* diagram for Example 5.

The speed within any interval of 1 second is a linear function and changes to another linear function in the next interval. The accumulated distance travelled by this robot in 10 seconds is equivalent to the area under this ascending zigzag curve. By treating each interval as a trapezoid, the total area under this zigzag curve is 30 units, or the exact accumulated distance travelled by this robot in 10 seconds is 30 m.

By applying the trapezium method to the *t*-*v* table above, the estimated total distance is 30 m, the same as the exact distance travelled. If using Sampson's method, the estimated total distance would be 28.3 m, with an error of 1.7 m. For this case, Simpson's method unnecessarily corrected the zigzag lines of any two adjacent intervals with a quadratic curve, which led to the error in the estimated distance. On the other hand, the trapezium method is a perfect fit to this case of 10 different-sized trapezoids.

## 3. Examples of incorporating alternative methods in numerical integration

The following three cases were reworked from classroom lectures or tutorials for the past engineering students at CQU to demonstrate how the numerical integration could be incorporated with other techniques to solve engineering and scientific problems. Example 6 was associated with the design of an underground network of pipelines. Example 7 was about approximate computing of the cardinal sine function sinc*x* that often occurs in engineering applications, such as communication, electronics, digital signal processing, and optical engineering. Example 8 was about numerical computing to estimate the accumulated distance that a particle travelled in the initial period. These problems have no analytical solutions and can only be approximated by numerical techniques with the pre-set error limits.

**Example 6**: In designing an underground network of pipelines, one section of the pipeline follows a cubic pattern $f(x)=\frac{1}{3}x^3$ in the local range 0–2 meters. Estimate the length of the pipeline in this section and be accurate to millimetres.

**Solution**

The length of any section of the function $f(x)=\frac{1}{3}x^3$ within [$a$, $b$] can be calculated by

$$L=\int_a^b\sqrt{1+(\frac{df}{dx})^2}\,dx=\int_a^b\sqrt{1+x^4}\,dx.$$

Plot of the integrand $y=\sqrt{1+x^4}$ in section [0, 2] is shown in Figure 7. Although this is a continuous and smooth function, no analytical solution to this integral can be found. Hence, numerical integration becomes the choice to estimate the length of pipeline in this section. Since the final estimate must be accurate to millimetres or error order of $10^{-3}$, the intermediate calculations should keep at least 4 decimal places or in the order of $10^{-4}$ to make sure that the final result is accurately rounded to the order of $10^{-3}$. Simpson's method should be chosen for this case as it is more accurate than the trapezium method under the same condition. If choosing the size of equal interval as $h$ = 0.1, the error bound for the composite Simpson's method should be around $O(h^4) = O(10^{-4})$. Therefore, the range [0, 2] must be divided into 20 equal intervals to ensure the required accuracy to millimetres. Following the similar process by Excel demonstrated in Examples 1-3 in the previous section, Simpson's method produced an estimated length of 3.653 m. If the trapezium method is used for the same division, the estimated length becomes 3.657 m. There would be a difference of 4 mm between the two estimates.

We can also divide the range into 18 equal intervals so that Simpson's 3/8 rule (6) can be used to check accuracy of the result from Simpson's method. This should produce an estimated length of 3.653 m, which is the same as the estimated length by Simpson's method using 20 equal intervals.

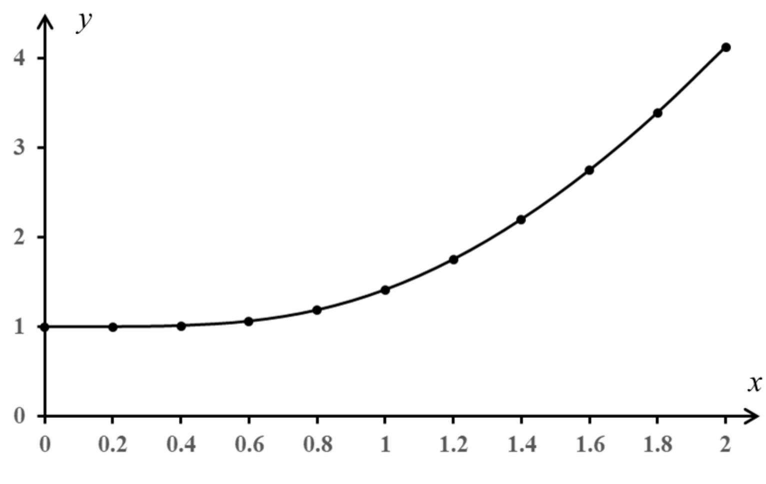


**Figure 7**. Curve of the integrand in Example 6 in section [0, 2].

**Example 7**: Approximate the integral of the cardinal sine function $\int_0^1 \frac{\sin x}{x} dx$ to be accurate to the order of $10^{-5}$.

**Solution**

Since this integral has no analytical solution, the challenging issue for this problem is not only about calculating an approximate value for the integral by Simpson's method with adequate number of subdivided strips, but also on how to verify that the value is met the required order of accuracy by other method. Because the accuracy is to the order of $10^{-5}$, or the error order of $O(10^{-6})$, the size of equal interval for Simpson's method should be around $h$ = 1/20 = 0.05 so that $h^4 = 0.05^4 = 6.26 \times 10^{-6}$. Following the similar process by Excel demonstrated in Examples 1-3 in the previous section, Simpson's method produced an estimated integral of 0.9460831, rounded down to 0.94608. If the trapezium method is used for the same division, the estimated integral becomes 0.94602. There would be a difference of 60 ppm (parts per million) between the two estimates. Note that in the numerical computing above, $y_0 = \lim_{x \to 0} \frac{\sin x}{x} = 1.$

We are confident that the result from Simpson's method for this case meets the required accuracy and likely has no error at least at the fifth decimal place under its error bound of $O(10^{-6})$. However, how accurate this estimate could be is still not validated. Maclaurin series can be used to obtain numerical solutions for some integrands that cannot be integrated by any integration techniques, such as $\int \frac{\sin x}{x} dx$. The Maclaurin series for sin$x$ is

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \frac{x^9}{9!} \ldots .$$

Hence,

$$\frac{\sin x}{x} = \frac{x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \frac{x^9}{9!} \ldots}{x} = 1 - \frac{x^2}{3!} + \frac{x^4}{5!} - \frac{x^6}{7!} + \frac{x^8}{9!} \ldots ;$$

$$\int_0^1 \frac{\sin x}{x} dx = \int_0^1 (1 - \frac{x^2}{3!} + \frac{x^4}{5!} - \frac{x^6}{7!} + \frac{x^8}{9!} \ldots) dx = (x - \frac{x^3}{3 \times 3!} + \frac{x^5}{5 \times 5!} - \frac{x^7}{7 \times 7!} + \frac{x^9}{9 \times 9!} \ldots) \Big|_0^1$$

$$= 1 - \frac{1}{3 \times 3!} + \frac{1}{5 \times 5!} - \frac{1}{7 \times 7!} + \frac{1}{9 \times 9!} \ldots$$

Since $\frac{1}{7 \times 7!} = 2.83 \times 10^{-5}$ and $\frac{1}{9 \times 9!} = 3.06 \times 10^{-7}$, the sum of the first five terms should make the error in the order of $10^{-7}$ or be accurate at least at the sixth decimal place that is

$$\int_0^1 \frac{\sin x}{x} dx \approx 1 - \frac{1}{3 \times 3!} + \frac{1}{5 \times 5!} - \frac{1}{7 \times 7!} + \frac{1}{9 \times 9!} = 0.9460831 .$$

This is the same as the result from Simpson's method with 20 equal intervals obtained above. Therefore, the error bound of $O(h^n)$ for Simpson's method is a conservative estimate on the error of approximation in common circumstances. The actual accuracy of estimation by Simpson's method is more likely to be higher than that indicated by the error bound of $O(h^n)$ in practice.

**Example 8**: A particle was accelerated initially. Five speed readings in metre per second (m/s) were recorded in the first five seconds as shown in the table below. Choose an appropriate method to estimate the accumulated distance that the particle travelled during this period. Use other means to verify the accuracy of the estimated distance. Keep 2 decimal places in the final result.

| *t* (second) | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| *v* (m/s) | 6.11 | 4.98 | 5.09 | 6.53 | 9.31 |

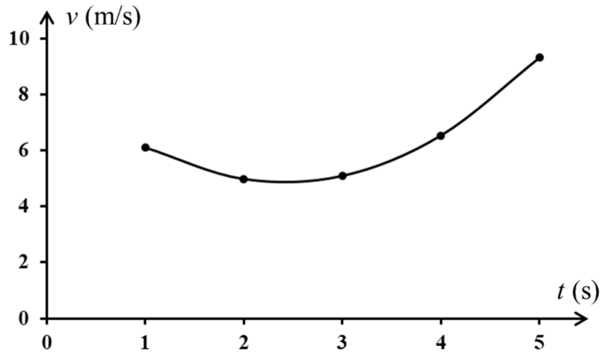


**Figure 8**. Plot of the speed readings in Example 8.

**Solution**

Plot of these speed readings in the first five seconds shows that the speed seems in a quadratic pattern during the period (Figure 8). Although the size of interval for this case is 1 second, the proportional time interval for estimating the error is $h$ = (single interval)/(total range) = 1s/4s = 0.25. If using Simpson's method, its error bound is around $h^4 = 0.25^4 = 3.9\times10^{-3}$, or in the order of $O(10^{-3})$. This error bound should be sufficient for the requirement of this problem to keep the final result with 2 decimal places. The trapezium method is not suitable for this case as its error bound is around $h^2 = 0.25^2 = 6.25\times10^{-2}$, or in the order of $O(10^{-2})$.

The distance travelled during the period is equivalent to the area under the curve connecting the speed readings in sequence in Figure 8. For the four intervals, Simpson's method produced an estimated distance of 23.88 m for the period, compared to that of 24.31 m produced by the trapezium method.

As this case is a sequence of discrete data, we are not able to use a known formula like Maclaurin series in Example 7 to verify the result from Simpson's method. However, we can use an 4$^{th}$-order interpolation formula by either the Lagrange or Newton's divided difference methods [4,5] to approximate this numerical integral. For example, by using Lagrange interpolation, a 4$^{th}$-order polynomial can be formulated by

$$L_4(t) = \sum_{i=1}^{5} v_i \prod_{j=1,\, j\neq i}^{5} \frac{t-t_j}{t_i-t_j}.$$

The process to get an interpolation used to be tedious by hand, but now one can easily get the formula from various software packages or advanced calculators. For example, by typing the five datasets in the table of speed readings into *WolframAlpha* [18], i.e., interpolation {(1,6.11), (2,4.98), (3,5.09), (4,6.53), (5,9.31)}, the following formula is displayed:

$$L_4(t) = -0.00333t^4 + 0.04833t^3 + 0.41333t^2 - 2.65833t + 8.31.$$

This formula is truncated at the fifth decimal place and hence accurate enough for this problem to be accurate to the second decimal place. Applying integration to this interpolation from first second to $5^{th}$ second should produce another estimate to the accumulated distance, i.e.,

$$\int_1^5 L_4(t)dt = \int_1^5 \left[-0.00333t^4 + 0.04833t^3 + 0.41333t^2 - 2.65833t + 8.31\right]dt \approx 23.8832.$$

This verifies that the accumulated distance estimated by Simpson's method is indeed accurate to centimetres, or the second decimal place.

## 4. Discussion and Conclusion

The eight examples demonstrated in this note shared some common concerns raised or mistakes made by engineering students at different stages of mathematics learning. Examples 1-3 outlined that the higher accuracy of approximation in numerical integration using Simpson's method may cause misconception to some students to take Simpson's method as the default choice to deal with numerical integration by overlooking the requirement that Simpson's method only applies to the cases with equal intervals of even numbers. Example 4 proven the consequence of misusing Simpson's method directly to nine data sets in an assessment item, which produced a result much worse than that by the trapezium method. Example 5 demonstrated that the trapezium method can overperform Simpson's method in special circumstances.

Example 6 demonstrated how to translate an engineering design to a problem involving numerical integration that does not have an analytical solution. By analysing the required accuracy, Simpson's method with an appropriate interval size was chosen to obtain the required outcome, which was then verified by Simpson's 3/8 rule. Example 7 demonstrated the same strategy of using Simpson's method to obtain the required integral for the cardinal sine function, but the accuracy was validated through Maclaurin series for the function. Example 8 demonstrated how to approach the accumulated distance confined under a set of sequential datasets and verify the accuracy of the estimated result by means of a Lagrange interpolation through available software or computing tools. All the examples solved by different numerical methods also led to the following observations.

- If an integrand is a well-defined continuous and smooth function in a given range, by adjusting the number of intervals (hence the width of the interval), Simpson's method would be able to produce an approximate solution satisfying the given accuracy or error limit. The error bound of $O(h^4)$ would hold true for most such cases and hence users should have confidence in the estimated error bound in practice.
- In cases where the integrand has a linear pattern between two adjacent known points (hence with a zigzag curve), the trapezium method would perform better than the Simpson's method. In cases where the integrand has high-frequency oscillations between any two adjacent known points, the trapezium method may also overperform Simpson's method [19].
- Under any circumstances, Simpson's method should not be directly applied to problems that have equal intervals of odd numbers as such would produce an estimate with a large error.

- If a set of discrete and sequential data points over a range is known, the estimated value to the integral over the range by applying any credible numerical method meeting the required accuracy should be regarded as a credible solution to the problem. This is because both the exact solution and the trend between any two adjacent points are unknown. Although different methods could be used for different cases, such as Simpson's 3/8 rule in Example 4 and a Lagrange interpolation in Example 8, they only provide new approximates, from which the estimate that is close to the majority of the approximated values may be regarded as the most appropriate result.
- In terms of learning and using numerical integration, choosing an approximate method from many practical options for a given problem is a task to ensure obtaining an acceptable solution according to the conditions. However, in cases where the exact solution is unknown, choosing another credible numerical method that can validate the obtained solution to the required accuracy or error limit is equally important.

This note only covered some common scenarios in using numerical integration to solve theoretical and practical problems. It would be much more helpful if a generalised framework could be established with accumulation of more studies in this area of research in the future. Also, these is a need to generalise Simpson's methods so that similar formulas could be applied to cases where equal intervals of odd numbers are only known, for which one alternative will be presented in the second part of this topic [20].

## Acknowledgments

All the past students whom the author taught in the applied calculus and advanced mathematics are appreciated for initiating and stimulating the works presented in this classroom note.

## Conflict of interest

The author declares no conflicts of interest in this paper.

## References

1. Guo, W.W., *Essentials and Examples of Applied Mathematics*, 2nd ed. 2020, Melbourne, Australia: Pearson.
2. Croft, A., Davison, R., Hargreaves, M. and Flint J., *Engineering Mathematics*, 5th ed. 2017, Harlow, UK: Pearson.
3. Guo, W., A practical strategy to improve performance of Newton's method in solving nonlinear equations. *STEM Education*, 2022, 2(4): 345–358. https://doi.org/10.3934/steme.2022021
4. Guo, W.W. and Wang, Y., *Advanced Mathematics for Engineering and Applied Sciences*, 2019, Sydney, Australia: Pearson.
5. Wang, Y. and Guo, W.W., *Applied Computational Modelling with MATLAB*, 2018, Melbourne: Pearson Australia.
6. Rozema, E., Estimating the error in the trapezoidal rule. *The American Mathematical Monthly*, 1980, 87(2): 124–128. https://doi.org/10.1080/00029890.1980.11994974

7. Cruz-Uribe, D. and Neugebauer, C.J., Sharp error bounds for the trapezoidal rule and Simpson's rule. *Journal of Inequalities in Pure and Applied Mathematics*, 2002, 3(4): Article 49.
8. Cruz-Uribe, D. and Neugebauer, C.J., An elementary proof of error estimates for the trapezoidal rule. *Mathematics Magazine*, 2003, 76(4): 303–306. https://doi.org/10.1080/0025570X.2003.11953199
9. Fazekas E.C. and Peter R. Mercer, P.R., Elementary proofs of error estimates for the midpoint and Simpson's rules. *Mathematics Magazine*, 2009, 82(5): 365–370. https://doi.org/10.4169/002557009X478418
10. Larson, R. and Edwards, B., *Calculus*, 12th ed. 2023, Boston, USA: Cengage.
11. Wheatley, G., *Applied Numerical Analysis*, 7th ed. 2004, Boston, USA: Pearson.
12. Chapra, S.C., *Applied Numerical Methods with MATLAB for Engineers and Scientists*, 2005, Boston, USA: McGraw-Hill Higher Education.
13. Ali, A.J. and Abbas, A.F., Applications of numerical integrations on the trapezoidal and Simpson's methods to analytical and MATLAB solutions. *Mathematical Modelling of Engineering Problems*, 2022, 9(5): 1352–1358. https://doi.org/10.18280/mmep.090525
14. Sauer, T., *Numerical Analysis*, 2nd ed. 2014, Harlow, UK: Pearson.
15. Guo, W., A guide for using integration by parts: *Pet-LoPo-InPo*. *Electronic Research Archive*, 2022, 30(10): 3572–3585. https://doi.org/10.3934/era.2022182
16. Guo, W., Streamlining applications of integration by parts in teaching applied calculus. *STEM Education*, 2022, 2(1): 73–83. https://doi.org/10.3934/steme.2022005
17. Guo, W. and Li, W., Simulating Vibrations of Two-Wheeled Self-balanced Robots with Road Excitations by MATLAB. In: Carbone, G., Laribi, M.A. (eds) Robot Design. *Mechanisms and Machine Science*, 2023, 123: 51–68. https://doi.org/10.1007/978-3-031-11128-0_3
18. WolframAlpha. Available from: https://www.wolframalpha.com/.
19. Kalambet, Y., Kozmin, Y. and Samokhin, A., Comparison of integration rules in the case of very narrow chromatographic peaks. *Chemometrics and Intelligent Laboratory Systems*. 2018, 179: 22–30. https://doi.org/10.1016/j.chemolab.2018.06.001
20. Guo, W., Solving problems involving numerical integration (II): Modified Simpson's method for general numeric integration. *STEM Education*, 2023, submitted for publication.


## Author's biography

**Dr. William Guo** is a professor in mathematics education with Central Queensland University, Australia. He is specialized in teaching applied mathematics for both engineering and education students. His research interests include mathematics education, applied computing, data analysis and numerical modelling. He is a member of IEEE and Australian Mathematical Society.