

A Comparison of Simpson's Rule Generalization through Lagrange and Hermite Interpolating Polynomials

Hasan Khanjar

College of Safety and Ocean Engineering, China University of Petroleum (Beijing), China

Email: hassankinjar@gmail.com

Received: 08 March, 2024; Revised: 20 April, 2024; Accepted: 26 May, 2024; Published: 08 September, 2024

Abstract: Simpson's Rule is a widely used numerical integration technique, but it cannot be applied to unequally spaced data. This paper presents a new generalization of Simpson's Rule using both Lagrange and Hermite interpolating polynomials to address this limitation. I provide a geometric interpretation of the method, showing its relationship to the area calculation of a trapezoid and a triangle, where the accuracy is significantly influenced by the chosen interpolating polynomial for midpoint determination. A comprehensive comparative analysis across various functions reveals that the Hermite-based approach consistently exhibits higher accuracy and stability than the Lagrange method, particularly with an increasing number of subintervals. This improved performance stems from the Hermite polynomial's ability to better approximate the function's behavior between data points. The findings highlight the effectiveness of the proposed Hermite-based generalization of Simpson's Rule in improving the accuracy of numerical integration for unequally spaced data, which is commonly encountered in practical applications.

Index Terms: Simpson's Rule, Numerical integration, unequally spaced data, accuracy improvement, Lagrange interpolating polynomial, Hermite interpolating polynomial, mathematical applications

1. Introduction

Numerical integration methods for unequally spaced data have been studied in various fields. Researchers have proposed formulas using techniques like Newton's Divided Difference formula to derive quadrature formulas, such as Trapezoidal and Simpson's rules, for unevenly spaced data [1, 2]. Additionally, in the context of time-series analysis, models like the complex irregular autoregressive model have been introduced to efficiently fit irregularly spaced time series data, allowing for accurate parameter estimation and forecasting of unobserved measurements [3]. Moreover, in spatial data analysis, a method based on Fourier transform has been described to approximate multiple integrals of covariance functions over irregular data regions with high accuracy and reduced computational cost [4]. These diverse approaches offer valuable tools for handling numerical integration with unequally spaced data. New interpolation, numerical differentiation, and numerical integration formulas with arbitrary order of accuracy are derived for evenly and unevenly spaced data [5].

Simpson's Rule is a widely used method for numerical integration, known for its accuracy and efficiency in approximating definite integrals. It is particularly well-suited for situations where the integrand is known only at a set of discrete points [6]. However, the accuracy of Simpson's Rule relies on the assumption of equally spaced data points, a condition that is often violated when dealing with real-world data [7,8]. This limitation arises due to various factors such as limitations in measurement techniques, the nature of the phenomena being studied, or constraints imposed by experimental setups.

Researchers have explored several approaches to adapt Simpson's Rule for unequally spaced data. One common strategy involves interpolating the given data to generate a new set of equally spaced points before applying Simpson's Rule [9]. Another approach involves modifying Simpson's Rule itself to directly account for the unequal spacing of data points [10]. This typically entails adjusting the weights assigned to each data point in the formula to reflect the varying distances between them. While research has explored modified Simpson's Rules and compared their accuracy and efficiency against other techniques like the Trapezoidal Rule and Gaussian quadrature [11,12], these modifications can often be complex and may not always yield significant accuracy improvements, particularly for functions with substantial curvature.

Despite these efforts, a robust and accurate method for applying Simpson's Rule to unequally spaced data, particularly for functions exhibiting significant curvature, remains a significant challenge in numerical analysis. To address this gap, this paper presents a new generalization of Simpson's Rule using both Lagrange and Hermite interpolating polynomials. These polynomial-based approaches provide a more flexible and potentially accurate way to handle unevenly spaced data while preserving the core principles of Simpson's Rule.

To provide a deeper understanding of my approach, I present a new geometric interpretation of Simpson's Rule, demonstrating its relationship to the area calculation of a trapezoid and a triangle. This interpretation reveals how the accuracy of the method is primarily linked to the chosen interpolation method for determining a crucial midpoint within each subinterval.

This paper conducts a rigorous comparative analysis of the accuracy and stability of Simpson's Rule generalized through both Lagrange and Hermite interpolating polynomials. By applying these methods to a diverse set of functions, including those exhibiting significant curvature, I demonstrate the strengths and limitations of each approach, particularly as the number of data points increases. Our analysis pays particular attention to error behavior, computational cost, and sensitivity to data distribution.

The findings of this research have important practical implications for any field that relies on accurate numerical integration of unequally spaced data. I discuss the potential benefits of my generalized Simpson's Rule methods in areas such as experimental data analysis, numerical simulation, and modeling of physical phenomena. Furthermore, I highlight the need for careful consideration of data characteristics and accuracy requirements when selecting the most appropriate interpolation method for a given application.

This research underscores the need for accurate and efficient numerical integration methods tailored for unequally spaced data, a challenge becoming increasingly prevalent in our data-driven world. The proposed Lagrange and Hermite-based generalizations of Simpson's Rule provide a valuable contribution to this effort, offering a balance of accuracy, efficiency, and applicability to a wide range of functions and scientific domains. Further exploration of adaptive algorithms holds promise for extending the impact of these methods to the increasingly large and complex datasets encountered in modern research and applications.

2. Preliminaries

2.1. Simpson's Rule

Let $a = x_0 < x_1 \dots < x_{2m} = b$, where $x_k = x_0 + kh$ and let $f_k = f(x_k)$, $0 \leq k \leq 2m$. Then, the generalized Simpson's rule is [13]

$$\int_a^b f(x)dx = \frac{h}{3}(f_0 + 2 \sum_{k=1}^{m-1} f_{2k} + 4 \sum_{k=1}^{m-1} f_{2k-1} + f_{2m}) \quad (1)$$

2.2. Lagrange Interpolating Polynomial

The Lagrange interpolation polynomial is a degree n polynomial that goes through $(n + 1)$ points $y_0 = f(x_0)$, $y_1 = f(x_1)$, $\dots$, $y_n = f(x_n)$. It can be expressed as follows [14]

$$P(x) = \sum_{j=0}^n P_j(x)$$

Where

$$P_j(x) = \sum_{j=0}^n y_j \prod_{k=0, k \neq j}^n \frac{x - x_k}{x_j - x_k}$$

2.3. Hermite Interpolating Polynomial

Theorem 1. If $f \in C^1[a, b]$ and $x_0, \dots, x_n \in [a, b]$ are distinct, the unique polynomial of least degree agreeing with f and f' at $x_0, \dots, x_n$ is the Hermite Polynomial of degree at most $2n + 1$ given by [15]

$$P(x) = \sum_{i=0}^n U_i(x)f(x_i) + \sum_{i=0}^n V_i(x)f'(x_i) \quad (2)$$

where $U_i(x)$ and $V_i(x)$ are polynomials in x of degree $(2n + 1)$, which can be obtained as follows

$$\begin{aligned} U_i(x_j) &= \begin{cases} 0 & \text{when } i \neq j; \\ 1 & \text{when } i = j \end{cases} \\ U'_i(x_j) &= 0 \text{ for all } i; \quad V_i(x_j) = \begin{cases} 0 & \text{when } i \neq j; \\ 1 & \text{when } i = j \end{cases} \end{aligned} \quad (3)$$

So, now we can say

$$U_i(x) = A_i(x)[L_i(x)]^2; \quad V_i(x) = B_i(x)[L_i(x)]^2$$

where

$$L_i(x) = \frac{(x - x_0)(x - x_1) \dots (x - x_n)}{(x_i - x_0)(x_i - x_1) \dots (x_i - x_n)}$$

Since the polynomial $[L_i(x)]^2$ has a degree of $2n$ and $U_i(x)$ and $V_i(x)$ have a degree of $(2n + 1)$, it follows that $A_i(x)$ and $B_i(x)$ are both linear equations. Therefore, we can write

$$\begin{aligned} U_i(x) &= (a_i + b_i x)[L_i(x)]^2 \\ V_i(x) &= (c_i + d_i x)[L_i(x)]^2 \end{aligned} \quad (4)$$

From (4) and (4), we obtain

$$\begin{aligned} a_i + b_i x_i &= 1, \quad c_i + d_i x_i = 0 \\ b_i + 2L'_i(x_i) &= 0, \quad d_i = 1 \end{aligned} \quad (5)$$

By solving equations in (5), we get

$$\begin{aligned} b_i &= -2L'_i(x_i), \quad a_i = 1 + 2x_i L'_i(x_i) \\ c_i &= -x_i, \quad d_i = 1 \end{aligned} \quad (6)$$

Substitute values in (6) in (4)

$$\begin{aligned} U_i(x) &= (1 - 2(x - x_i)L'_i(x_i))[L_i(x)]^2 \\ V_i(x) &= (x - x_i)[L_i(x)]^2 \end{aligned} \quad (7)$$

Finally, by substituting $U_i(x)$ and $V_i(x)$ into (2), we get

$$P(x) = \sum_{i=0}^n (1 - 2(x - x_i)L'_i(x_i))[L_i(x)]^2 f(x_i) + \sum_{i=0}^n (x - x_i)[L_i(x)]^2 f'(x_i)$$

3. Main Results

Proposition 1. For any three points (x_0, y_0) , (x_1, y_1) , and (x_2, y_2) the area (A) of the triangle formed by these points is given by the formula

$$A = \frac{1}{2} [x_0(y_2 - y_1) + x_1(y_0 - y_2) + x_2(y_1 - y_0)]$$

Proof. Consider a triangle ΔABC in the Cartesian plane, where $A(x_0, y_0)$, $B(x_1, y_1)$, and $C(x_2, y_2)$ are its vertices. Let N, M , and L be the x-axis points, and O is the origin.

Then $NM = OM - ON = x_1 - x_0$, $NL = OL - ON = x_2 - x_0$, and $ML = OL - OM = x_2 - x_1$. Now, we express the area A as the sum of the areas of these trapeziums

$$\begin{aligned} A &= \frac{1}{2} [(AM + BN) \times NM + (AM + CL) \times ML - (BN + CL) \times NL] \\ &= \frac{1}{2} [(y_1 + y_0)(x_1 - x_0) + (y_1 + y_2)(x_2 - x_1) - (y_0 + y_2)(x_2 - x_0)] \\ &= \frac{1}{2} [x_0(y_2 - y_1) + x_1(y_0 - y_2) + x_2(y_1 - y_0)] \end{aligned}$$

Theorem 2. Let $P(x)$ be a polynomial with a degree of three or lower that is defined on the interval $[a, b]$. Let A , B , and C be different points on the curve $y = P(x)$. Denote the midpoint of the interval by C . If A_t represents the total area enclosed by the curve $y = P(x)$ and the line segments AC and BC , then it holds that

$$A_t = \frac{1}{3} \cdot \text{Area}(\Delta ABC) \quad (8)$$

And the area under the curve

$$\int_a^b f(x) dx = \frac{b-a}{2} [f(a) + f(b)] + \frac{4}{3} A \quad (9)$$

Where A is ΔABC area

Proof. Consider a polynomial $P(x)$ of degree three or lower defined on the interval $[a, b]$. Let $A = (x_1, f(x_1))$, $B = (x_2, f(x_2))$, and $C = \left(\frac{x_1+x_2}{2}, f\left(\frac{x_1+x_2}{2}\right)\right)$ be three different points on the curve $y = P(x)$. The total area over the interval $[a, b]$ is

$$\begin{aligned} A_t &= \int_{x_1}^{x_2} f(x) dx = \int_{x_1}^{x_2} [a_0 x^2 + a_1 x + a_2] dx = \left[\frac{a_0 x^3}{3} + \frac{a_1 x^2}{2} + a_2 x \right]_{x_1}^{x_2} \\ &= \left(\frac{x_2^3 - x_1^3}{3} \right) a_0 + \left(\frac{x_2^2 - x_1^2}{2} \right) a_1 + a_2 (x_2 - x_1) \end{aligned} \quad (10)$$

We get the area of ΔABC from proposition 1

$$\begin{aligned} A &= \frac{1}{2} \left[\frac{x_1 + x_2}{2} (f(x_1) - f(x_2)) + x_1 \left(f(x_2) - f\left(\frac{x_1 + x_2}{2}\right) \right) + x_2 \left(f\left(\frac{x_1 + x_2}{2}\right) - f(x_1) \right) \right] \\ &= \frac{1}{2} \left[\left[\frac{x_1 + x_2}{2} - x_2 \right] (a_0 x_1^2 + a_1 x_1 + a_2) + \left[x_1 - \frac{x_1 + x_2}{2} \right] (a_0 x_2^2 + a_1 x_2 + a_2) \right. \\ &\quad \left. + [x_2 - x_1] \left(a_0 \left(\frac{x_1 + x_2}{2} \right)^2 + a_1 \left(\frac{x_1 + x_2}{2} \right) + a_2 \right) \right] \\ &= \frac{1}{2} \left[\left[\frac{x_1 - x_2}{2} \right] (a_0 x_1^2 + a_1 x_1 + a_2) + \left[\frac{x_1 - x_2}{2} \right] (a_0 x_2^2 + a_1 x_2 + a_2) \right. \\ &\quad \left. + [x_2 - x_1] \left(a_0 \left(\frac{x_1 + x_2}{2} \right)^2 + a_1 \left(\frac{x_1 + x_2}{2} \right) + a_2 \right) \right] \\ &= \frac{1}{4} \left[(x_1 - x_2) (a_0 x_1^2 + a_1 x_1 + a_2) + (x_1 - x_2) (a_0 x_2^2 + a_1 x_2 + a_2) \right. \\ &\quad \left. + 2[x_2 - x_1] \left(a_0 \left(\frac{x_1 + x_2}{2} \right)^2 + a_1 \left(\frac{x_1 + x_2}{2} \right) + a_2 \right) \right] \\ &= \frac{1}{4} \left[(x_1 - x_2) (a_0 x_1^2 + a_1 x_1 + a_2) + (x_1 - x_2) (a_0 x_2^2 + a_1 x_2 + a_2) \right. \\ &\quad \left. + [x_2 - x_1] \left(\frac{a_0}{2} (x_1^2 + 2x_1 x_2 + x_2^2) + a_1 (x_1 + x_2) + 2a_2 \right) \right] \\ &= \frac{1}{4} \left[a_0 \left[(x_1 - x_2) x_1^2 + (x_1 - x_2) x_2^2 + \frac{x_1^2 + 2x_1 x_2 + x_2^2}{2} (x_2 - x_1) \right] \right. \\ &\quad \left. + a_1 \left[\frac{(x_1 - x_2) x_1 + (x_1 - x_2) x_2}{(x_2 - x_1)(x_1 + x_2)} \right] \right. \\ &\quad \left. + a_2 [(x_1 - x_2) + (x_1 - x_2) + 2(x_2 - x_1)] \right] \\ &= \frac{1}{4} \left[a_0 \left[x_1^3 - x_2 x_1^2 + x_2^2 x_1 - x_2^3 + \frac{x_2 x_1^2 + 2x_1 x_2^2 + x_2^3 - x_1^3 - 2x_1^2 x_2 - x_1 x_2^2}{2} \right] \right. \\ &\quad \left. + a_1 \left[\frac{x_1^2 - x_1 x_2}{x_1 x_2 - x_2^2 + x_1 x_2 + x_2^2 - x_1^2 - x_1 x_2} \right] \right. \\ &\quad \left. + a_2 [x_1 - x_2 + x_1 - x_2 + 2x_2 - 2x_1] \right] \\ &= \frac{1}{4} \left[\left[\frac{2x_1^3 - 2x_2 x_1^2 + 2x_2^2 x_1 - 2x_2^3 + x_2 x_1^2 + 2x_1 x_2^2 + x_2^3 - x_1^3 - 2x_1^2 x_2 - x_1 x_2^2}{2} \right] a_0 \right] \\ &= \left[\frac{x_1^3 - x_2^3 - 3x_2 x_1^2 + 3x_1 x_2^2}{8} \right] a_0 \end{aligned} \quad (11)$$

Area of trapezoid will be

$$\begin{aligned} & \frac{x_2 - x_1}{2} (a_0 x_1^2 + a_1 x_1 + a_2 + a_0 x_2^2 + a_1 x_2 + a_2) \\ &= \left(\frac{x_2^3 - x_1^3 + x_2 x_1^2 - x_1 x_2^2}{2} \right) a_0 + \left(\frac{x_2^2 - x_1^2}{2} \right) a_1 + (x_2 - x_1) a_2 \end{aligned} \quad (12)$$

Subtract (10) from (11) and (12)

$$\begin{aligned} & \left(\frac{x_2^3 - x_1^3}{3} \right) a_0 + \left(\frac{x_2^2 - x_1^2}{2} \right) a_1 + a_2 (x_2 - x_1) - a_0 \left[\frac{x_1^3 - x_2^3 - 3x_2 x_1^2 + 3x_1 x_2^2}{8} \right] \\ & \quad - \left(\frac{x_2^3 - x_1^3 + x_2 x_1^2 - x_1 x_2^2}{2} \right) a_0 - \left(\frac{x_2^2 - x_1^2}{2} \right) a_1 - (x_2 - x_1) a_2 \\ & \left[\frac{x_2^3 - x_1^3}{3} - \frac{x_1^3 - x_2^3 - 3x_2 x_1^2 + 3x_1 x_2^2}{8} - \frac{x_2^3 - x_1^3 + x_2 x_1^2 - x_1 x_2^2}{2} \right] a_0 + \left[\frac{x_2^2 - x_1^2}{2} - \frac{x_2^2 - x_1^2}{2} \right] a_1 \\ & \quad \left[\frac{8x_2^3 - 8x_1^3 - 3x_1^3 + 3x_2^3 + 9x_2 x_1^2 - 9x_1 x_2^2 - 12x_2^3 + 12x_1^3 - 12x_2 x_1^2 + 12x_1 x_2^2}{24} \right] a_0 \\ & \quad = \left[\frac{x_1^3 - x_2^3 - 3x_2 x_1^2 + 3x_1 x_2^2}{24} \right] a_0 \end{aligned} \quad (13)$$

We can see that (13) is 1/3 of (11)

Corollary 1. Since Simpson's 1/3 Rule is derived from a quadratic polynomial, we can use Theorem 2 with it.

Proof. select three successive equally spaced points, denoted as x_{j-1} , x_j , and x_{j+1} , as following

$$\begin{aligned} &= \frac{1}{2} (x_{j+1} - x_{j-1}) [f(x_{j-1}) + f(x_{j+1})] \\ & \quad + \frac{2}{3} [x_j (f(x_{j-1}) - f(x_{j+1})) + (x_{j-1}) (f(x_{j+1}) - f(x_j)) + x_{j+1} (f(x_j) - f(x_{j-1}))] \\ &= x_j \left[\frac{2}{3} f(x_{j-1}) - \frac{2}{3} f(x_{j+1}) \right] + x_{j-1} \left[-\frac{1}{2} f(x_{j-1}) - \frac{1}{2} f(x_{j+1}) + \frac{2}{3} f(x_{j+1}) - \frac{2}{3} f(x_j) \right] \\ & \quad + x_{j+1} \left[\frac{1}{2} f(x_{j-1}) + \frac{1}{2} f(x_{j+1}) + \frac{2}{3} f(x_j) - \frac{2}{3} f(x_{j-1}) \right] \\ &= x_j \left[\frac{2}{3} f(x_{j-1}) - \frac{2}{3} f(x_{j+1}) \right] + x_{j-1} \left[-\frac{1}{2} f(x_{j-1}) + \frac{1}{6} f(x_{j+1}) - \frac{2}{3} f(x_j) \right] + x_{j+1} \left[\frac{1}{2} f(x_{j+1}) + \frac{2}{3} f(x_j) - \frac{1}{6} f(x_{j-1}) \right] \\ &= \frac{1}{6} x_j [4f(x_{j-1}) - 4f(x_{j+1})] + \frac{1}{6} x_{j-1} [-3f(x_{j-1}) + f(x_{j+1}) - 4f(x_j)] + \frac{1}{6} x_{j+1} [3f(x_{j+1}) + 4f(x_j) - f(x_{j-1})] \\ &= \frac{1}{6} f(x_{j-1}) [4x_j - 3x_{j-1} - x_{j+1}] + \frac{1}{6} f(x_j) [-4x_{j-1} + 4x_{j+1}] + \frac{1}{6} f(x_{j+1}) [-4x_j + x_{j-1} + 3x_{j+1}] \\ &= \frac{1}{6} [f(x_{j-1}) [4x_j - 3x_{j-1} - x_{j+1}] + f(x_j) [-4x_{j-1} + 4x_{j+1}] + f(x_{j+1}) [-4x_j + x_{j-1} + 3x_{j+1}]] \\ &= \frac{1}{6} [f(x_{j-1}) [4x_j - 3x_{j-1} - x_{j+1}] + f(x_j) [-4x_{j-1} + 4x_{j+1}] + f(x_{j+1}) [-4x_j + x_{j-1} + 3x_{j+1}]] \end{aligned}$$

Considering that $x_{j-1} = x_j - h$ and $x_{j+1} = x_j + h$, we obtain

$$\begin{aligned} &= \frac{1}{6} [f(x_{j-1}) [4h + 4x_j - 4h + 3h - 3x_j - x_j - h] + f(x_j) [-4x_j + 4h + 4x_j + 4h] \\ & \quad + f(x_{j+1}) [-4x_{j-1} - 4h + x_j - h + 3x_j + 3h]] \\ &= \frac{1}{6} [f(x_{j-1}) [2h] + f(x_j) [8h] + f(x_{j+1}) [2h]] \\ &= \frac{h}{3} [f(x_{j-1}) + 4f(x_j) + f(x_{j+1})] \end{aligned} \quad (14)$$

We can notice that (14) is identical to (1)

4. Simpson's Rule for Unequally spaced data

The error in Simpson's Rules when applied to unequally spaced data comes from the distortion of the triangle involved. The triangle $\triangle ABC$ should have C as its midpoint, which does not hold true for unequally spaced data because C will not be at the middle distance in the interval of integration. To address this problem, we can provide a midpoint for the triangle by utilizing different interpolating polynomials techniques.

4.1 Lagrange interpolating polynomial

To provide a midpoint for the triangle by utilizing Lagrange interpolating polynomial, let x_0 , x_1 , and x_2 represent three consecutive, unequally spaced points, and let $x = \frac{x_0+x_2}{2}$. Lagrange Interpolating Polynomial at x will be

$$f(x) = \frac{(x-x_1)(x-x_2)}{(x_0-x_1)(x_0-x_2)}f(x_0) + \frac{(x-x_0)(x-x_2)}{(x_1-x_0)(x_1-x_2)}f(x_1) + \frac{(x-x_0)(x-x_1)}{(x_2-x_0)(x_2-x_1)}f(x_2) \quad (15)$$

Using proposition 1, we can get the area of $\triangle ABC$

$$\begin{aligned} A &= \frac{1}{2} [x(f(x_0) - f(x_2)) + x_0(f(x_2) - f(x)) + x_2(f(x) - f(x_0))] \\ 2A &= xf(x_0) - xf(x_2) + x_0f(x_2) - x_0f(x) + x_2f(x) - x_2f(x_0) \\ 2A &= [x - x_2]f(x_0) + [x_0 - x]f(x_2) + [x_2 - x_0]f(x) \end{aligned} \quad (16)$$

Substitute (15) in (16)

$$\begin{aligned} 2A &= [x - x_2]f(x_0) + [x_0 - x]f(x_2) + \frac{(x - x_1)(x - x_2)}{(x_1 - x_0)}f(x_0) + \frac{(x_2 - x_0)(x - x_0)(x - x_2)}{(x_1 - x_0)(x_1 - x_2)}f(x_1) \\ &\quad + \frac{(x - x_0)(x - x_1)}{(x_2 - x_1)}f(x_2) \\ 2A &= \left[(x - x_2) + \frac{(x - x_1)(x - x_2)}{(x_1 - x_0)} \right] f(x_0) + \left[(x_0 - x) + \frac{(x - x_0)(x - x_1)}{(x_2 - x_1)} \right] f(x_2) \\ &\quad + \frac{(x_2 - x_0)(x - x_0)(x - x_2)}{(x_1 - x_0)(x_1 - x_2)} f(x_1) \\ 2A &= \left[\left(\frac{x_0 - x_2}{2} \right) + \frac{(x_0 + x_2 - 2x_1)(x_0 - x_2)}{4(x_1 - x_0)} \right] f(x_0) + \left[\left(\frac{x_0 - x_2}{2} \right) + \frac{(x_2 - x_0)(x_0 + x_2 - 2x_1)}{4(x_2 - x_1)} \right] f(x_2) \\ &\quad + \frac{(x_2 - x_0)^3}{4(x_1 - x_0)(x_2 - x_1)} f(x_1) \\ 2A &= \left[\left(\frac{x_0 - x_2}{2} \right) + \frac{(x_0 + x_2 - 2x_1)(x_0 - x_2)}{4(x_1 - x_0)} \right] f(x_0) + \left[\left(\frac{x_0 - x_2}{2} \right) + \frac{(x_2 - x_0)(x_0 + x_2 - 2x_1)}{4(x_2 - x_1)} \right] f(x_2) \\ &\quad + \frac{(x_2 - x_0)^3}{4(x_1 - x_0)(x_2 - x_1)} f(x_1) \end{aligned} \quad (17)$$

Substitute (17) in (9)

$$\begin{aligned} &\frac{x_2 - x_0}{2} [f(x_0) + f(x_2)] \\ &\quad + \frac{2}{3} \left[\left[\left(\frac{x_0 - x_2}{2} \right) + \frac{(x_0 + x_2 - 2x_1)(x_0 - x_2)}{4(x_1 - x_0)} \right] f(x_0) \right. \\ &\quad \left. + \left[\left(\frac{x_0 - x_2}{2} \right) + \frac{(x_2 - x_0)(x_0 + x_2 - 2x_1)}{4(x_2 - x_1)} \right] f(x_2) + \frac{(x_2 - x_0)^3}{4(x_1 - x_0)(x_2 - x_1)} f(x_1) \right] \\ &= \frac{1}{3} \left[\left[x_0 - x_2 + \frac{(x_0 + x_2 - 2x_1)(x_0 - x_2)}{2(x_1 - x_0)} + \frac{3}{2}(x_2 - x_0) \right] f(x_0) \right. \\ &\quad \left. + \left[x_0 - x_2 + \frac{(x_2 - x_0)(x_0 + x_2 - 2x_1)}{2(x_2 - x_1)} + \frac{3}{2}(x_2 - x_0) \right] f(x_2) + \frac{(x_2 - x_0)^3}{2(x_1 - x_0)(x_2 - x_1)} f(x_1) \right] \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{3} \left[\left(\frac{x_2 - x_0}{2} \right) + \frac{(x_0 + x_2 - 2x_1)(x_0 - x_2)}{2(x_1 - x_0)} \right] f(x_0) + \left[\left(\frac{x_2 - x_0}{2} \right) + \frac{(x_2 - x_0)(x_0 + x_2 - 2x_1)}{2(x_2 - x_1)} \right] f(x_2) \\
&\quad + \frac{(x_2 - x_0)^3}{2(x_1 - x_0)(x_2 - x_1)} f(x_1) \Bigg] \\
&= \frac{x_2 - x_0}{6} \left[\left[1 + \frac{(x_0 + x_2 - 2x_1)}{(x_0 - x_1)} \right] f(x_0) + \left[1 + \frac{(x_0 + x_2 - 2x_1)}{(x_2 - x_1)} \right] f(x_2) + \frac{(x_2 - x_0)^2}{(x_1 - x_0)(x_2 - x_1)} f(x_1) \right] \\
&= \frac{x_2 - x_0}{6} \left[\left[\frac{x_0 + x_2 - 2x_1 + x_0 - x_1}{(x_0 - x_1)} \right] f(x_0) + \left[\frac{x_0 + x_2 - 2x_1 + x_2 - x_1}{(x_2 - x_1)} \right] f(x_2) + \frac{(x_2 - x_0)^2}{(x_1 - x_0)(x_2 - x_1)} f(x_1) \right] \\
&= \frac{x_2 - x_0}{6} \left[\left[\frac{2x_0 + x_2 - 3x_1}{(x_0 - x_1)} \right] f(x_0) + \frac{(x_2 - x_0)^2}{(x_1 - x_0)(x_2 - x_1)} f(x_1) + \left[\frac{x_0 + 2x_2 - 3x_1}{(x_2 - x_1)} \right] f(x_2) \right]
\end{aligned} \tag{18}$$

Equation (18) transforms into conventional Simpson's Rule if the step size is equally spaced. Therefore, this formula can be applied to both equally and unequally spaced points.

The Composite Simpson's rule can be obtained as follows

$$\begin{aligned}
&\sum_{i=0}^{n/2-1} \frac{x_{2i+2} - x_{2i}}{6} \left[\left[\frac{2x_{2i} + x_{2i+2} - 3x_{2i+1}}{(x_{2i} - x_{2i+1})} \right] f(x_{2i}) + \frac{(x_{2i+2} - x_{2i})^2}{(x_{2i+1} - x_{2i})(x_{2i+2} - x_{2i+1})} f(x_{2i+1}) \right. \\
&\quad \left. + \left[\frac{x_{2i} + 2x_{2i+2} - 3x_{2i+1}}{(x_{2i+2} - x_{2i+1})} \right] f(x_{2i+2}) \right]
\end{aligned} \tag{19}$$

4.2 Hermite interpolating polynomial

To provide a midpoint for the triangle through Hermite Interpolating Polynomial, let x_0 , x_1 , and x_2 represent three consecutive, unequally spaced points, and let $x = \frac{x_0 + x_2}{2}$. Hermite Interpolating Polynomial at x will be

$$\begin{aligned}
H_5 \left(\frac{x_2 + x_0}{2} \right) &= u_0 \left(\frac{x_2 + x_0}{2} \right) \cdot f(x_0) + v_0 \left(\frac{x_2 + x_0}{2} \right) \cdot f'(x_0) + u_1 \left(\frac{x_2 + x_0}{2} \right) \cdot f(x_1) + v_1 \left(\frac{x_2 + x_0}{2} \right) \cdot f'(x_1) \\
&\quad + u_2 \left(\frac{x_2 + x_0}{2} \right) \cdot f(x_2) + v_2 \left(\frac{x_2 + x_0}{2} \right) \cdot f'(x_2)
\end{aligned}$$

We can get u_i and v_i for $i = 0, 1, 2$ from (7)

$$\begin{aligned}
H_5 \left(\frac{x_2 + x_0}{2} \right) &= \frac{(3x_0 - 2x_1 - x_2)(x_2 + x_0 - 2x_1)^2}{16(x_0 - x_1)^3} \cdot f(x_0) + \left(\frac{x_2 - x_0}{32} \right) \left[\frac{x_2 + x_0 - 2x_1}{(x_0 - x_1)} \right]^2 \cdot f'(x_0) \\
&\quad + \left[1 - \frac{(x_2 + x_0 - 2x_1)^2}{(x_0 - x_1)(x_1 - x_2)} \right] \frac{(x_2 - x_0)^4}{16(x_1 - x_0)^2(x_1 - x_2)^2} \cdot f(x_1) \\
&\quad + \left(\frac{x_2 + x_0 - 2x_1}{32} \right) \left[\frac{(x_2 - x_0)^2}{(x_1 - x_0)(x_1 - x_2)} \right]^2 \cdot f'(x_1) + \left[1 + \frac{2x_2 - x_0 - x_1}{(x_2 - x_1)} \right] \frac{(x_2 + x_0 - 2x_1)^2}{16(x_2 - x_1)^2} f(x_2) \\
&\quad + \left(\frac{x_0 - x_2}{32} \right) \left[\frac{x_2 + x_0 - 2x_1}{(x_2 - x_1)} \right]^2 \cdot f'(x_2)
\end{aligned}$$

By providing the midpoint obtained from Hermite to the triangle in (9) we get

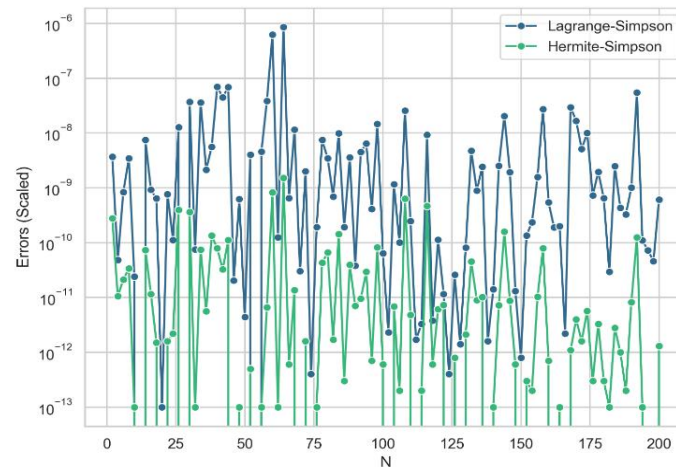
$$\begin{aligned}
& \frac{x_2 - x_0}{2} [f(x_0) + f(x_2)] \\
& + \frac{2}{3} \left[[x_2 - x_0] \left[\left(\frac{(3x_0 - 2x_1 - x_2)(x_2 + x_0 - 2x_1)^2}{16(x_0 - x_1)^3} - \frac{1}{2} \right) \cdot f(x_0) \right. \right. \\
& + \left(\frac{x_2 - x_0}{32} \right) \left[\frac{x_2 + x_0 - 2x_1}{(x_0 - x_1)} \right]^2 \cdot f'(x_0) + \left[1 - \frac{(x_2 + x_0 - 2x_1)^2}{(x_0 - x_1)(x_1 - x_2)} \right] \frac{(x_2 - x_0)^4}{16(x_1 - x_0)^2(x_1 - x_2)^2} \cdot f(x_1) \\
& + \left(\frac{x_2 + x_0 - 2x_1}{32} \right) \left[\frac{(x_2 - x_0)^2}{(x_1 - x_0)(x_1 - x_2)} \right]^2 \cdot f'(x_1) \\
& \left. + \left(\left[1 + \frac{2x_2 - x_0 - x_1}{(x_2 - x_1)} \right] \frac{(x_2 + x_0 - 2x_1)^2}{16(x_2 - x_1)^2} - \frac{1}{2} \right) f(x_2) + \left(\frac{x_0 - x_2}{32} \right) \left[\frac{x_2 + x_0 - 2x_1}{(x_2 - x_1)} \right]^2 \cdot f'(x_2) \right] \Bigg] \\
& = \frac{(x_2 - x_0)}{6} \left[\left(\frac{(3x_0 - 2x_1 - x_2)(x_2 + x_0 - 2x_1)^2}{4(x_0 - x_1)^3} + 1 \right) \cdot f(x_0) + \left(\frac{x_2 - x_0}{8} \right) \left[\frac{x_2 + x_0 - 2x_1}{(x_0 - x_1)} \right]^2 \cdot f'(x_0) \right. \\
& + \left[1 - \frac{(x_2 + x_0 - 2x_1)^2}{(x_0 - x_1)(x_1 - x_2)} \right] \frac{(x_2 - x_0)^4}{4(x_1 - x_0)^2(x_1 - x_2)^2} \cdot f(x_1) \\
& + \left(\frac{x_2 + x_0 - 2x_1}{8} \right) \left[\frac{(x_2 - x_0)^2}{(x_1 - x_0)(x_1 - x_2)} \right]^2 \cdot f'(x_1) \\
& \left. + \left(\left[1 + \frac{2x_2 - x_0 - x_1}{(x_2 - x_1)} \right] \frac{(x_2 + x_0 - 2x_1)^2}{4(x_2 - x_1)^2} + 1 \right) f(x_2) + \left(\frac{x_0 - x_2}{8} \right) \left[\frac{x_2 + x_0 - 2x_1}{(x_2 - x_1)} \right]^2 \cdot f'(x_2) \right]
\end{aligned}$$

And the composite form will be as follows

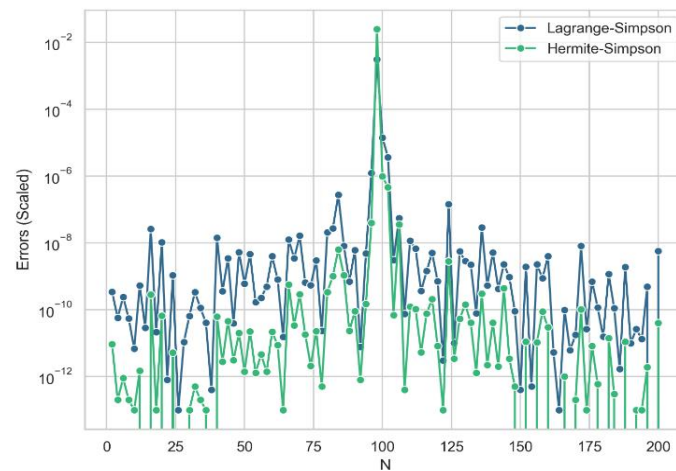
$$\begin{aligned}
& \sum_{i=0}^{n/2-1} \frac{(x_{2i+2} - x_{2i})}{6} \left[\left(\frac{(3x_{2i} - 2x_{2i+1} - x_{2i+2})(x_{2i+2} + x_{2i} - 2x_{2i+1})^2}{4(x_{2i} - x_{2i+1})^3} + 1 \right) \cdot f(x_{2i}) \right. \\
& + \left(\frac{x_{2i+2} - x_{2i}}{8} \right) \left[\frac{x_{2i+2} + x_{2i} - 2x_{2i+1}}{(x_{2i} - x_{2i+1})} \right]^2 \cdot f'(x_{2i}) \\
& + \left[1 - \frac{(x_{2i+2} + x_{2i} - 2x_{2i+1})^2}{(x_{2i} - x_{2i+1})(x_{2i+1} - x_{2i+2})} \right] \frac{(x_{2i+2} - x_{2i})^4}{4(x_{2i+1} - x_{2i})^2(x_{2i+1} - x_{2i+2})^2} \cdot f(x_{2i+1}) \\
& + \left(\frac{x_{2i+2} + x_{2i} - 2x_{2i+1}}{8} \right) \left[\frac{(x_{2i+2} - x_{2i})^2}{(x_{2i+1} - x_{2i})(x_{2i+1} - x_{2i+2})} \right]^2 \cdot f'(x_{2i+1}) \\
& + \left(\left[1 + \frac{2x_{2i+2} - x_{2i} - x_{2i+1}}{(x_{2i+2} - x_{2i+1})} \right] \frac{(x_{2i+2} + x_{2i} - 2x_{2i+1})^2}{4(x_{2i+2} - x_{2i+1})^2} + 1 \right) f(x_{2i+2}) \\
& \left. + \left(\frac{x_{2i} - x_{2i+2}}{8} \right) \left[\frac{x_{2i+2} + x_{2i} - 2x_{2i+1}}{(x_{2i+2} - x_{2i+1})} \right]^2 \cdot f'(x_{2i+2}) \right] \quad (20)
\end{aligned}$$

5. Numerical Results and Discussion

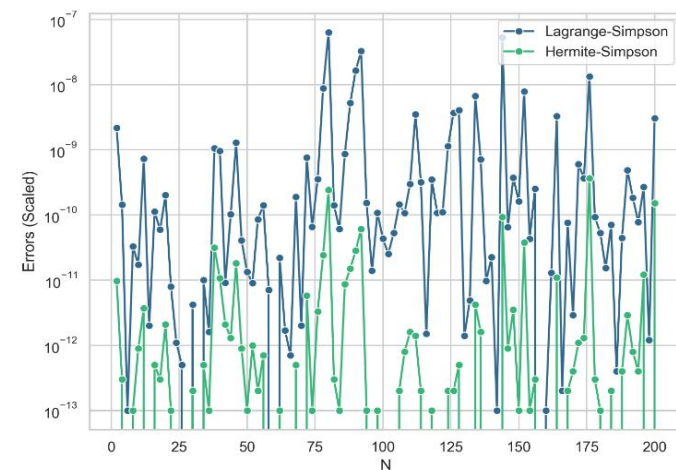
Real-life applications of numerical analysis functions will be examined over subintervals $n = 200$, and then we will compare the error produced by Simpson's Rule generalized from Lagrange and Hermite interpolating polynomials (see Fig. 1).



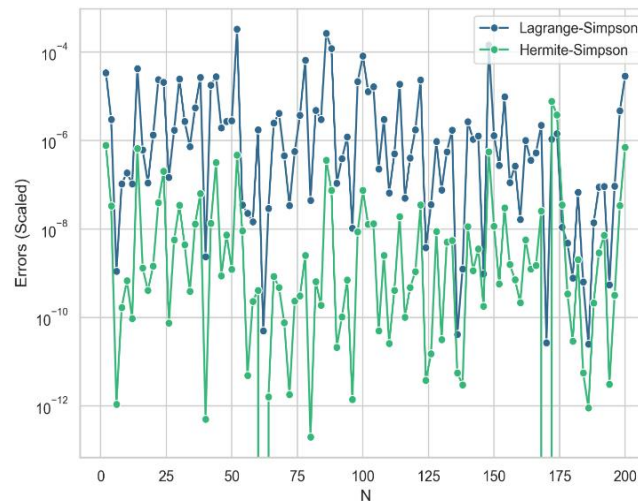
(a) $\text{erf}(2)$



(b) The volume flow rate $\int_0^R 5 \left(1 - \frac{x}{R}\right)^{1/7} 2\pi x \, dx$ in a smooth circular pipe with radius $R = 0.2$



(c) The normal distribution $\frac{1}{\sigma\sqrt{2\pi}} \int_0^u e^{-\frac{(x-u)^2}{2\sigma^2}} dx$ with standard deviation $\sigma = 5$ and a mean $u = 10$.



- (d) The moment of normal forces $\frac{aPabr}{\sin(\theta_a)} \int_{\theta_1}^{\theta_2} \sin^2(\theta) d\theta$ on a Block-Type Hand Brake with a maximum pressure on the shoe $P_a = 27.4 \text{ psi}$, face width $b = 1.25 \text{ in}$, radius $r = 6 \text{ in}$, $a = 10$, $\theta_1 = 8.13^\circ$, $\theta_2 = 98.13^\circ$, and $\theta_a = \theta_1 - \theta_2$

Fig. 1. The figures (a) through (d) show error analysis for various functions using generalized Simpson from Lagrange and Hermite methods. The graph's y-axis is scaled using a logarithmic scale to present the data effectively.

Fig. 1 shows that in various functions, generalized Simpson from Hermite exhibits lower error rates. However, this is not universally applicable. In uncommon cases, Lagrange shows lower error than Hermite (see Fig. 2). This anomaly can be attributed to the error margins of Lagrange method. This level of inaccuracy in certain scenarios, can paradoxically lead to more accurate results when dealing with higher-order polynomials (e.g., $x^4, x^5, \dots, x^n$) or complex functions. The error introduced by Lagrange's method may manifest as a positive error. Incorporating this error into the calculated area may enhance the overall accuracy of the method, resulting in a more precise estimate.

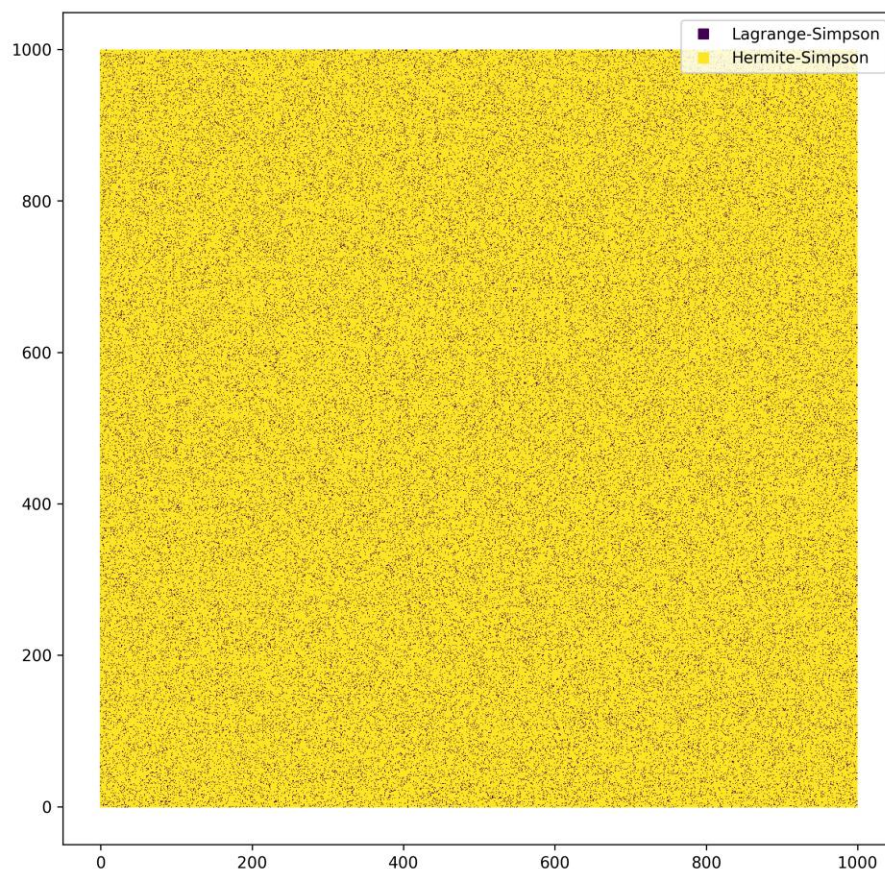


Fig. 2. The probability of Hermite Simpson and Lagrange Simpson generalizations outperforming each other across one million mesh grids.

The Hermite method demonstrates a high level of stability when compared with Lagrange in generalizing Simpson's Rule. The location of c within the interval $[a, b]$, where a, b , and c are three unevenly distributed points, has almost no significant impact on its performance. This does not hold for the Lagrange method, as the position of c significantly influences the resulting error, which decreases as we approach the midpoint of the interval (at midpoint $c = 1$, Hermite and Lagrange produce equivalent errors). Fig. 3 shows this behavior.

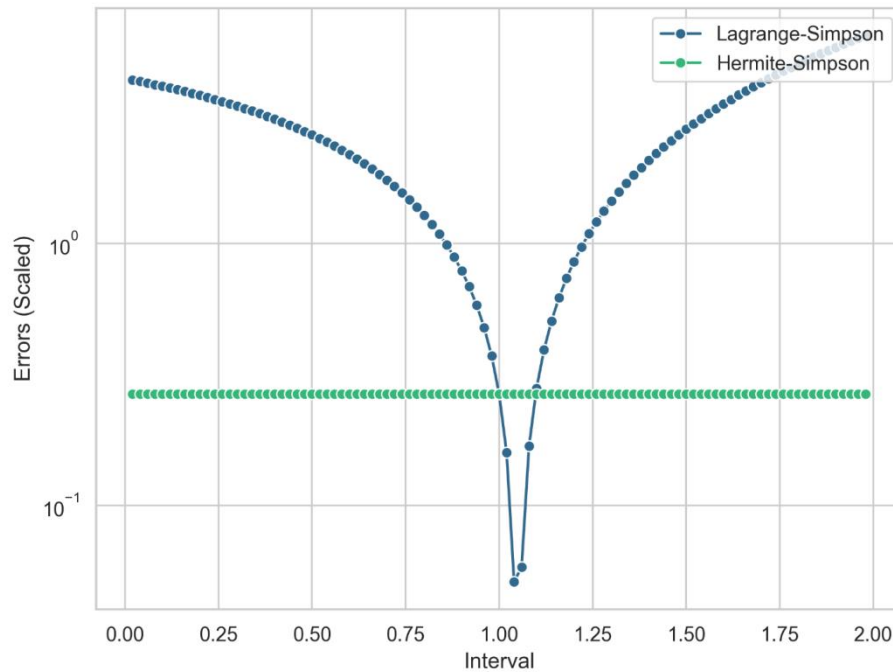
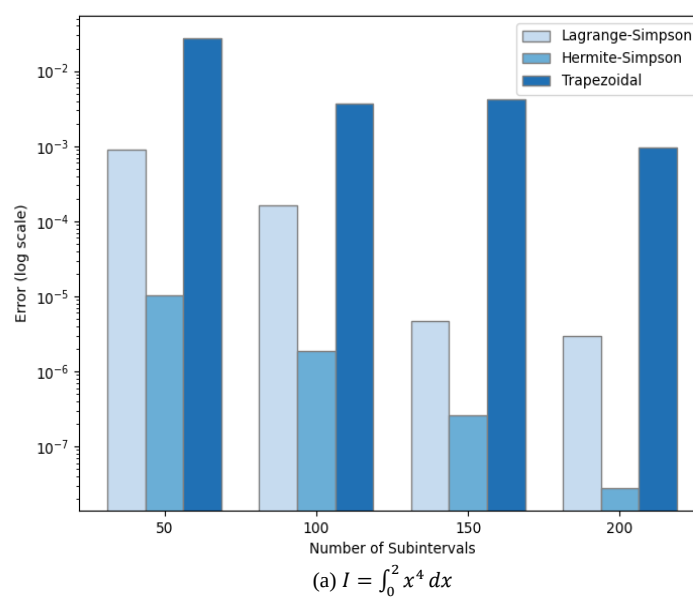
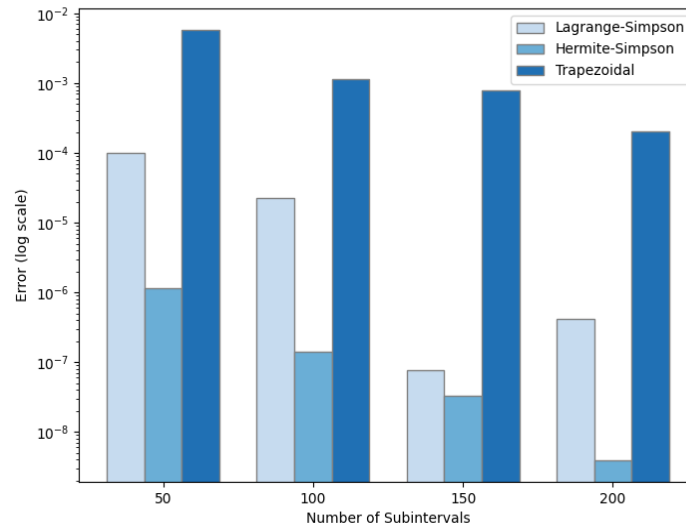


Fig. 3. Error analysis for the integration $\int_0^2 x^4 dx$ with varying values of c

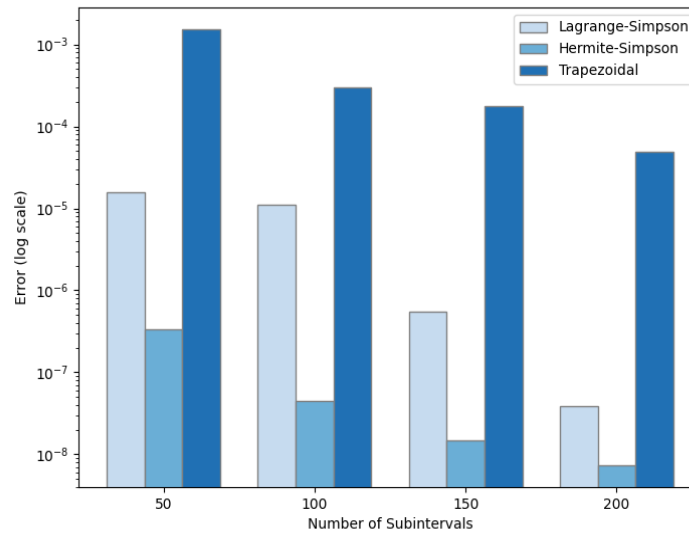
In the example given, 198 different values for ' c ' were utilized. Out of these, in 193 cases, Hermite outperformed Lagrange (97.5%), while Lagrange performed better in 4 cases (2%). Finally, in one case, their performances were found to be equal (0.05%).

Furthermore, I employed the Trapezoidal Rule as a reference point for assessing the effectiveness of the proposed methods, given its potential adaptability to unequally spaced data. Fig. 4 shows the error analysis for each method across various functions for different subintervals (50, 100, 150, 200).

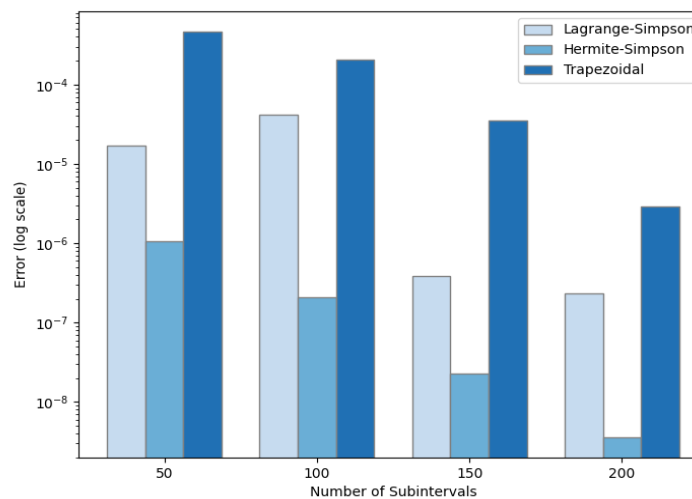




$$(b) I = \int_0^2 e^x dx$$



$$(c) I = \int_0^2 \sin(x) dx$$



$$(d) I = \int_0^2 \frac{1}{1+x^2} dx$$

Fig. 4. Performance comparison of Lagrange-Simpson, Hermite-Simpson, and Trapezoidal numerical integration methods.

Fig. 4 shows that the Hermite-Simpson method emerges as the most accurate and reliable integration method across all tested functions and subintervals. The Trapezoidal method, while simpler, exhibits the highest error terms, making it less suitable for high-accuracy requirements.

In general, the error term decreases with an increasing number of subintervals. However, the Lagrange-Simpson method presents an interesting pattern in Fig. 4b and d., where the error term increases when $n = 200$ for b and $n = 100$ for d.— indicating that this method depends on the distribution of the data used. This trend cannot be seen in Trapezoidal and Hermite-Simpson methods which both show stability.

The superior performance of the Hermite-based Simpson's Rule can be attributed to the fact that Hermite interpolation incorporates both function values and derivatives, providing a more faithful representation of the function's behavior within subintervals, particularly for functions with high curvature. While computationally more demanding, the increased accuracy offered by the Hermite generalization could be crucial in applications where even small integration errors can significantly impact results, such as in high-precision engineering design or financial modeling.

6. Conclusions

This research presents a new generalization of Simpson's Rule, addressing the challenge of accurately integrating functions from unequally spaced data. By employing Lagrange and Hermite interpolating polynomials, I enhance the accuracy and applicability of this widely used numerical integration technique. My findings conclusively demonstrate that the Hermite-based approach consistently surpasses the Lagrange method in terms of accuracy and stability, particularly with an increasing number of data points. However, it is more computationally costly than Lagrange because it takes into account both the function and derivative values. Although there are some cases where Lagrange outperforms Hermite, this holds true in only about 10% of cases. This anomaly can be attributed to the error margins of Lagrange method, which in some cases paradoxically lead to more accurate results. The proposed Hermite-based generalization of Simpson's Rule offers a robust and accurate solution for integrating a wide range of functions from unevenly spaced data, a problem frequently encountered in real-world applications. The rigorous comparative analysis performed, encompassing error behavior, computational cost, and sensitivity to data distribution, further strengthens the validity and practical relevance of our findings.

The enhanced accuracy achieved through the Hermite-based method translates directly into practical benefits across diverse fields reliant on precise numerical integration of unequally spaced data. For instance, in experimental data analysis, my method enables more reliable analysis of measurements with inherent uneven spacing, common due to instrument limitations or experimental constraints. This improvement is particularly valuable in applications such as calculating work done by variable forces, estimating heat transfer in systems with fluctuating temperatures, and predicting satellite trajectories from irregularly timed observations. Future research should explore the development of adaptive algorithms based on the specific characteristics of the data. Finally, applying these new methods to multi-dimensional integration problems with unequally spaced data presents a promising avenue for addressing increasingly complex scientific challenges.

Limitations

The study did not cover the analysis of the error terms for Generalized Simpson's Rule using Hermite and Lagrange interpolating polynomials. Additionally, although Theorem 2 is applicable for Simpson 3/8 Rule, there was no derivation for it.

References

- [1] Md., Nayan, Dhali., Nandita, Barman., Md., Mohedul, Hasan., A., K., M., Selim, Reza. (2020). Numerical Double Integration for Unequal Data Spaces. *International Journal of Mathematical Sciences and Computing*, doi: 10.5815/IJMSC.2020.06.04
- [2] Md., Mamun-Ur-Rashid, Khan., Mohammad, Rubaiyat, Tanvir, Hossain., Selina, Parvin. (2017). Numerical Integration Schemes for Unequal Data Spacing. *American Journal of Applied Mathematics*, doi: 10.11648/J.AJAM.20170502.12
- [3] Felipe, Elorrieta., Felipe, Elorrieta., Susana, Eyheramendy., Wilfredo, Palma., Wilfredo, Palma. (2019). Discrete-time autoregressive model for unequally spaced time-series observations. *Astronomy and Astrophysics*, doi: 10.1051/0004-6361/201935560
- [4] Peter, Simonson., Douglas, Nychka., Soutir, Bandyopadhyay. (2020). Rapid Numerical Approximation Method for Integrated Covariance Functions Over Irregular Data Regions. doi: 10.1002/STA4.275
- [5] Kumar, M. R. (n.d.). New Formulas and Methods for Interpolation, Numerical Differentiation and Numerical Integration.
- [6] Bjorck, A., Bajpai, A. C., Calus, I. M., & Fairley, J. A. (1978). *Numerical Methods for Engineers and Scientists*. Mathematics of Computation, 32(141). <https://doi.org/10.2307/2006283>
- [7] Burden, R. L., & Faires, J. D. (2011). *Numerical Analysis 9th Edition*. In Brooks/Cole (Vol. 4, Issue 3).
- [8] E. Ward Cheney and David R. Kincaid. 2007. *Numerical Mathematics and Computing* (6th. ed.). Brooks/Cole Publishing Co., USA.
- [9] Cameron Reed, B. (2014). Numerically integrating irregularly-spaced (x, y) data. *Mathematics Enthusiast*, 11(3). <https://doi.org/10.54870/1551-3440.1319>
- [10] Isaac, A., Boakye Stephen, T., & Seidu, B. (2021). A New Trapezoidal-Simpson 3/8 Method for Solving Systems of Nonlinear Equations. *American Journal of Mathematical and Computer Modelling*, 6(1). <https://doi.org/10.11648/j.ajmcm.20210601.11>

- [11] Mamun-Ur-Rashid Khan, Md. (2017). Numerical Integration Schemes for Unequal Data Spacing. American Journal of Applied Mathematics, 5(2). <https://doi.org/10.11648/j.ajam.20170502.12>
- [12] NUGA, O. A. (2019). Comparison on Trapezoidal and Simpson's Rule for Unequal Data Space. International Journal of Mathematical Sciences and Computing, 5(4), 24–32. <https://doi.org/10.5815/ijmsc.2019.04.03>
- [13] Rogers, A. D. (2007). Integrals of Fitted Polynomials and an Application to Simpson's Rule. The College Mathematics Journal, 38(2). <https://doi.org/10.1080/07468342.2007.11922227>
- [14] Udovičić, Z. (2006). SOME MODIFICATIONS OF THE TRAPEZOIDAL RULE. Sarajevo Journal of Mathematics, 2(15).
- [15] Woodford, C., & Phillips, C. (2012). Numerical Integration. In Numerical Methods with Worked Examples: Matlab Edition. https://doi.org/10.1007/978-94-007-1366-6_5

Author's Profiles



Hasan Khanjar currently pursuing a Master's degree at China University of Petroleum (Beijing). I obtained my Bachelor's degree in Petroleum Engineering from Baghdad University, where I studied from 2016 to 2020. My research interests encompass gas hydrate, numerical simulation, hydraulic fracturing, and machine learning.

I was awarded the prestigious CSC scholarship from China. My dedication to research, especially in areas that demand significant intellectual effort, drives my academic journey. My ORCID is 0000-0002-4936-4409. For academic inquiries or professional connections.

How to cite this paper: Hasan Khanjar, "A Comparison of Simpson's Rule Generalization through Lagrange and Hermite Interpolating Polynomials", International Journal of Mathematical Sciences and Computing(IJMSC), Vol.10, No.3, pp. 37-50, 2024. DOI: 10.5815/ijmsc.2024.03.04