

# Numerical Integration

The notes for this week have been obtained from > 1. Scheid, F. J. (1968). Schaum's outline of theory and problems of numerical analysis. > 2. Chasnov, J. R. (2012). Numerical methods. Hong Kong University of Science and Technology. <https://www.math.hkust.edu.hk/~machas/numerical-methods.pdf>

## The trapezoidal rule

$$\int_{x_0}^{x_n} y(x) dx \approx \frac{1}{2}h(y_0 + 2y_1 + \cdots + 2y_{n-1} + y_n)$$

is an elementary, but typical, **composite numerical integration formula**, known as the **composite trapezoidal rule**. This rule approximates the area under the curve  $y(x)$  over the interval  $[x_0, x_n]$  by dividing the interval into  $n$  equal subintervals of width  $h = \frac{x_n - x_0}{n}$ , and using **straight line segments** to connect successive points  $(x_i, y_i)$ . These segments form trapezoids, whose areas are straightforward to compute.

The formula effectively replaces the integrand  $y(x)$  with a piecewise linear function, interpolating  $y(x)$  at the mesh points  $x_0, x_1, \dots, x_n$ . The sum of the areas of these trapezoids gives the approximation to the integral.

This approach is simple yet remarkably effective for sufficiently smooth functions. However, like any numerical method, it introduces an error, known as the **truncation error**, which quantifies the deviation from the exact integral. For the trapezoidal rule, the error is approximately

$$-\frac{(x_n - x_0)h^2}{12}y^{(2)}(\xi)$$

for some  $\xi \in [x_0, x_n]$ , where  $y^{(2)}(\xi)$  denotes the second derivative of  $y(x)$  at some point  $\xi$  in the integration interval. This expression shows that the error decreases quadratically with the step size  $h$ , assuming  $y^{(2)}(x)$  remains bounded. Consequently, halving the step size reduces the error by roughly a factor of four, reflecting the **second-order accuracy** of the method.

Thus, the trapezoidal rule provides a good balance between simplicity and accuracy, particularly for functions that are at least twice continuously differentiable on the interval of integration.

### Solved Problem 14.5

Apply the trapezoidal rule to approximate the integral of  $\sqrt{x}$  over the interval  $[1.00, 1.30]$  using six equally spaced subintervals. Then compare the result with the exact value of the integral.

Dividing the interval  $[1.00, 1.30]$  into  $n = 6$  subintervals gives a step size of  $h = 0.05$ . The corresponding partition points are:

$$x_i = 1.00, 1.05, 1.10, 1.15, 1.20, 1.25, 1.30$$

Evaluating the function  $y(x) = \sqrt{x}$  at these points yields:

$$y_i = 1.00000, 1.02470, 1.04881, 1.07238, 1.09545, 1.11803, 1.14017$$

Applying the trapezoidal rule:

$$\begin{aligned} \int_{1.00}^{1.30} \sqrt{x} \, dx &\approx \frac{0.05}{2} [y_0 + 2(y_1 + y_2 + y_3 + y_4 + y_5) + y_6] \\ &= \frac{0.05}{2} [1.00000 + 2(1.02470 + 1.04881 + 1.07238 + 1.09545 + 1.11803) + 1.14017] \\ &= \frac{0.05}{2} [1.00000 + 2(5.35937) + 1.14017] = \frac{0.05}{2} \cdot 12.85891 = 0.32147 \end{aligned}$$

The exact value of the integral is:

$$\int_{1.00}^{1.30} \sqrt{x} \, dx = \frac{2}{3} \left( 1.3^{3/2} - 1 \right) \approx 0.32149$$

Thus, the trapezoidal approximation is accurate to five decimal places, with an absolute error of approximately 0.00002.

### Solved Problem 14.6

Derive an estimate of the truncation error of the trapezoidal rule.

To estimate the truncation error of the composite trapezoidal rule, consider approximating the definite integral

$$\int_a^b y(x) \, dx$$

by dividing the interval  $[a, b]$  into  $n$  equal subintervals of width  $h = \frac{b-a}{n}$ . On each subinterval  $[x_i, x_{i+1}]$ , the trapezoidal rule approximates the area under  $y(x)$  by using a linear segment between  $y(x_i)$  and  $y(x_{i+1})$ , forming a trapezoid. The total error in the approximation comes from the accumulation of local errors over each subinterval.

**Local Truncation Error via Taylor Expansion** Consider a single subinterval  $[x_i, x_{i+1}]$ , and denote  $x = x_i + t$ , where  $0 \leq t \leq h$ . Let's expand  $y(x)$  about  $x_i$  using Taylor's theorem:

$$y(x_i + t) = y(x_i) + ty'(x_i) + \frac{t^2}{2}y''(\xi_t),$$

for some  $\xi_t \in [x_i, x_i + t] \subseteq [x_i, x_{i+1}]$ .

Now integrate  $y(x)$  over the interval:

$$\int_{x_i}^{x_{i+1}} y(x) dx = \int_0^h \left[ y(x_i) + ty'(x_i) + \frac{t^2}{2}y''(\xi_t) \right] dt.$$

Compute each term separately:

- $\int_0^h y(x_i) dt = hy(x_i)$
- $\int_0^h ty'(x_i) dt = \frac{h^2}{2}y'(x_i)$
- $\int_0^h \frac{t^2}{2}y''(\xi_t) dt = \frac{1}{2} \int_0^h t^2 y''(\xi_t) dt$

Since  $y''(\xi_t)$  varies with  $t$ , we can't integrate it exactly, but by the **mean value theorem for integrals**, there exists some point  $\xi_i \in [x_i, x_{i+1}]$  such that:

$$\int_0^h t^2 y''(\xi_t) dt = y''(\xi_i) \int_0^h t^2 dt = y''(\xi_i) \cdot \frac{h^3}{3}.$$

Thus, the full integral becomes:

$$\int_{x_i}^{x_{i+1}} y(x) dx = hy(x_i) + \frac{h^2}{2}y'(x_i) + \frac{1}{2} \cdot \frac{h^3}{3}y''(\xi_i) = hy(x_i) + \frac{h^2}{2}y'(x_i) + \frac{h^3}{6}y''(\xi_i).$$

Now consider the **trapezoidal approximation** over the same interval:

$$\frac{h}{2} [y(x_i) + y(x_{i+1})].$$

We expand  $y(x_{i+1})$  using Taylor's theorem again:

$$y(x_{i+1}) = y(x_i + h) = y(x_i) + hy'(x_i) + \frac{h^2}{2}y''(\tilde{\xi}_i)$$

for some  $\tilde{\xi}_i \in [x_i, x_{i+1}]$ . Substituting into the trapezoidal approximation:

$$\frac{h}{2} \left[ y(x_i) + y(x_i) + hy'(x_i) + \frac{h^2}{2} y''(\tilde{\xi}_i) \right] = hy(x_i) + \frac{h^2}{2} y'(x_i) + \frac{h^3}{4} y''(\tilde{\xi}_i).$$

Now subtract the trapezoidal approximation from the exact integral:

$$\begin{aligned} E_i &= \int_{x_i}^{x_{i+1}} y(x) dx - \frac{h}{2} [y(x_i) + y(x_{i+1})] \\ &= \left( hy(x_i) + \frac{h^2}{2} y'(x_i) + \frac{h^3}{6} y''(\xi_i) \right) - \left( hy(x_i) + \frac{h^2}{2} y'(x_i) + \frac{h^3}{4} y''(\tilde{\xi}_i) \right) \\ &= \frac{h^3}{6} y''(\xi_i) - \frac{h^3}{4} y''(\tilde{\xi}_i) = -\frac{h^3}{12} y''(\bar{\xi}_i) \end{aligned}$$

for some  $\bar{\xi}_i \in [x_i, x_{i+1}]$ , assuming  $y''$  is smooth enough to interpolate between  $\xi_i$  and  $\tilde{\xi}_i$ .

**Global Truncation Error** Summing the local errors over all subintervals gives the total truncation error:

$$E = \sum_{i=0}^{n-1} E_i = -\frac{h^3}{12} \sum_{i=0}^{n-1} y''(\bar{\xi}_i).$$

Assuming  $y''(x)$  is continuous, the sum can be approximated as  $ny''(\xi)$  for some  $\xi \in [a, b]$ . Substituting  $nh = b - a$ , we obtain the classical global error estimate:

$$\text{Truncation error} = -\frac{(b-a)h^2}{12} y''(\xi),$$

where  $\xi \in [a, b]$ . This result shows that the global error of the trapezoidal rule is proportional to  $h^2$  and depends on the second derivative of the integrand.

### Solved Problem 14.7

Apply the estimate of Problem 14.6 to our square root integral.

To estimate the error in the trapezoidal approximation of the integral

$$\int_{1.00}^{1.30} \sqrt{x} dx,$$

we apply the global truncation error formula:

$$\text{Error} \approx -\frac{(b-a)h^2}{12} y^{(2)}(\xi),$$

where  $h = 0.05$ ,  $b - a = 0.30$ , and the second derivative of  $y(x) = \sqrt{x}$  is

$$y^{(2)}(x) = -\frac{1}{4}x^{-3/2}.$$

Using  $\xi = 1.0$  to obtain a conservative estimate, we compute:

$$\text{Error} \approx -\frac{0.30 \cdot (0.05)^2}{12} \cdot \left(-\frac{1}{4}\right) = 0.000015625.$$

This estimated error is slightly less than the actual error of 0.00002, as previously observed. However, when the trapezoidal approximation 0.32147 is corrected by adding this estimated error, the result becomes

$$0.32147 + 0.000015625 = 0.32149,$$

which matches the exact value of the integral to five decimal places. Thus, the error estimate is consistent and provides a reliable correction to the numerical approximation.

### Solved Problem 14.8

Estimate the effect of inaccuracies in the  $y_k$  values on results obtained by the trapezoidal rule.

With  $Y_k$  denoting the true values, as before, we find

$$\frac{h}{2}(e_0 + 2e_1 + \cdots + 2e_{n-1} + e_n)$$

as the error introduced into the trapezoidal approximation due to inaccuracies in the  $y_k$  values, where  $e_k = Y_k - y_k$  represents the deviation of each computed or measured value  $y_k$  from the true value  $Y_k$ .

This expression arises because the trapezoidal rule weights the endpoint values  $y_0$  and  $y_n$  once, and all interior values  $y_1, \dots, y_{n-1}$  twice. Consequently, the total contribution of the errors  $e_k$  to the integral follows the same weighting structure.

Now, suppose the magnitude of each error is bounded by some constant  $E$ , so that

$$|e_k| \leq E \quad \text{for all } k = 0, 1, \dots, n.$$

In the worst case, all errors reach this upper bound and contribute maximally to the output error. Therefore, we can bound the total error as follows:

$$\left| \frac{h}{2}(e_0 + 2e_1 + \cdots + 2e_{n-1} + e_n) \right| \leq \frac{h}{2}(E + 2(n-1)E + E).$$

Simplifying:

$$= \frac{h}{2} \cdot 2nE = nhE.$$

Since  $nh = b - a$ , this gives the final bound:

$$\text{Maximum output error} \leq (b - a)E.$$

This result shows that the overall impact of uniformly bounded errors in the  $y_k$  values is linearly proportional to both the length of the integration interval and the maximum pointwise error. It offers a useful and intuitive measure of how sensitive the trapezoidal rule is to inaccuracies in the input data.

#### Solved Problem 14.9

Apply the above to the square root integral of Problem 14.5.

We have  $(b - a)E = (0.30)(0.00005) = 0.000015$ , so that this source of error is negligible.

#### Simpson's rule

$$\int_{x_0}^{x_n} y(x) dx \approx \frac{h}{3}(y_0 + 4y_1 + 2y_2 + 4y_3 + \cdots + 2y_{n-2} + 4y_{n-1} + y_n)$$

is also a composite formula, and comes from using connected parabolic segments as the approximation to  $y(x)$ . It is one of the most heavily used formulas for approximate integration. The truncation error is about

$$-(x_n - x_0) \frac{h^4}{180} y^{(4)}(\xi)$$

for some  $\xi \in [x_0, x_n]$ .

#### Solved Problem 14.10

Derive the composite Simpson's Rule by approximating the integrand with quadratic interpolants over subintervals and assembling a general integration formula.

**Step 1: The Setup** We begin with the task of approximating the definite integral:

$$\int_a^b f(x) dx$$

Suppose the interval  $[a, b]$  is subdivided into  $n$  equal subintervals, where  $n$  is **even**. Define:

- Step size:

$$h = \frac{b - a}{n}$$

- Nodes:

$$x_k = a + kh, \quad \text{for } k = 0, 1, \dots, n$$

- Function values:

$$y_k = f(x_k)$$

We'll derive Simpson's Rule by approximating the integrand  $f(x)$  over **pairs of subintervals** using a **quadratic polynomial** through three consecutive points.

**Step 2: Approximating over a Single Segment** Consider three consecutive points:

$$x_0, \quad x_1 = x_0 + h, \quad x_2 = x_0 + 2h$$

Using **Lagrange interpolation**, we construct a quadratic polynomial  $P(x)$  that satisfies:

$$P(x_k) = y_k = f(x_k), \quad \text{for } k = 0, 1, 2$$

The interpolating polynomial is:

$$P(x) = y_0 \frac{(x - x_1)(x - x_2)}{(x_0 - x_1)(x_0 - x_2)} + y_1 \frac{(x - x_0)(x - x_2)}{(x_1 - x_0)(x_1 - x_2)} + y_2 \frac{(x - x_0)(x - x_1)}{(x_2 - x_0)(x_2 - x_1)}$$

Integrating  $P(x)$  from  $x_0$  to  $x_2$ , we find:

$$\int_{x_0}^{x_2} P(x) dx = \frac{h}{3}(y_0 + 4y_1 + y_2)$$

This is **Simpson's Rule over a single pair of intervals**.

**Step 3: Generalizing to the Full Interval** Now, we apply this idea repeatedly across the entire interval  $[a, b]$ , grouping points in sets of three:

- From  $x_0$  to  $x_2$ :

$$\frac{h}{3}(y_0 + 4y_1 + y_2)$$

- From  $x_2$  to  $x_4$ :

$$\frac{h}{3}(y_2 + 4y_3 + y_4)$$

- From  $x_4$  to  $x_6$ :

$$\frac{h}{3}(y_4 + 4y_5 + y_6)$$

- ...

- Up to  $x_{n-2}$  to  $x_n$ :

$$\frac{h}{3}(y_{n-2} + 4y_{n-1} + y_n)$$

Adding these expressions, the middle terms appear multiple times and must be combined. The full expression becomes:

$$\int_a^b f(x) dx \approx \frac{h}{3} (y_0 + 4y_1 + 2y_2 + 4y_3 + 2y_4 + \cdots + 2y_{n-2} + 4y_{n-1} + y_n)$$

### Solved Problem 14.11

Apply Simpson's rule to the integral of Problem 14.5.

We apply Simpson's Rule to approximate the integral:

$$\int_{1.00}^{1.30} \sqrt{x} dx$$

using 6 equally spaced nodes with step size  $h = 0.05$ . The Simpson's Rule formula gives:

$$\int_{1.00}^{1.30} \sqrt{x} dx \approx \frac{0.05}{3} [1.0000 + 4(1.02470 + 1.07238 + 1.11803) + 2(1.04881 + 1.09544) + 1.14017]$$

Evaluating the expression:

$$= \frac{0.05}{3} [1.0000 + 4(3.21511) + 2(2.14425) + 1.14017] = \frac{0.05}{3} (19.28911) = 0.32149$$

which is correct to five places.



### Solved Problem 14.12

Estimate the truncation error of Simpson's Rule.

To understand the truncation error of Simpson's Rule, we begin by examining what happens over a single application of the rule — that is, over just two subintervals.

Suppose we have three points:  $x_0$ ,  $x_1 = x_0 + h$ , and  $x_2 = x_0 + 2h$ . On the interval  $[x_0, x_2]$ , we construct a quadratic interpolating polynomial  $P(x)$  that matches the function  $f(x)$  at those three points. Simpson's Rule then approximates the integral  $\int_{x_0}^{x_2} f(x) dx$  by integrating the polynomial  $P(x)$  instead:

$$\int_{x_0}^{x_2} f(x) dx \approx \int_{x_0}^{x_2} P(x) dx = \frac{h}{3}(f(x_0) + 4f(x_1) + f(x_2)).$$

This approximation is exact if  $f(x)$  is a polynomial of degree at most three. However, for a general smooth function, there is a difference between the actual function and its quadratic interpolant. This difference is captured by an error term from interpolation theory, which says that

$$f(x) - P(x) = \frac{(x - x_0)(x - x_1)(x - x_2)}{3!} f^{(3)}(\xi_x)$$

for some  $\xi_x \in [x_0, x_2]$ . When this expression is integrated over  $[x_0, x_2]$ , the result is a precise formula for the **local truncation error** of Simpson's Rule:

$$E_{\text{local}} = \int_{x_0}^{x_2} f(x) dx - \int_{x_0}^{x_2} P(x) dx = -\frac{h^5}{90} f^{(4)}(\xi),$$

for some  $\xi \in [x_0, x_2]$ . This tells us that the approximation deviates from the true integral by a term involving the fourth derivative of the function and the fifth power of the step size. The power of  $h^5$  reflects how rapidly the local error shrinks as the intervals get finer.

Now, if the full interval  $[a, b]$  is divided into  $n$  equal subintervals with even  $n$ , then Simpson's Rule is applied  $n/2$  times — once per pair of subintervals. Each application contributes its own local error term of the form above, and the total error is the sum of these:

$$E = -\frac{h^5}{90} \left( f^{(4)}(\xi_1) + f^{(4)}(\xi_2) + \cdots + f^{(4)}(\xi_{n/2}) \right).$$

Assuming  $f^{(4)}$  is continuous, this sum can be approximated as

$$E \approx -\frac{h^5}{90} \cdot \frac{n}{2} f^{(4)}(\xi)$$

for some  $\xi \in [a, b]$ . Substituting  $nh = b - a$ , the total truncation error becomes

$$\boxed{\text{Truncation error} = -\frac{(b-a)h^4}{180}f^{(4)}(\xi), \quad \xi \in [a, b].}$$

This final result shows that Simpson's Rule has fourth-order accuracy, with the error shrinking proportionally to  $h^4$  and depending on how curved the function is, as measured by the fourth derivative.

### Solved Problem 14.13

Apply the estimate of Problem 14.12 to our square root integral.

We apply the Simpson's Rule truncation error formula to the integral

$$\int_{1.00}^{1.30} \sqrt{x} \, dx,$$

which was evaluated using 6 subintervals of width  $h = 0.05$ . The truncation error formula from Problem 14.12 is

$$\text{Error} = -\frac{(b-a)h^4}{180}f^{(4)}(\xi), \quad \text{for some } \xi \in [a, b].$$

For  $f(x) = \sqrt{x} = x^{1/2}$ , the fourth derivative is

$$f^{(4)}(x) = -\frac{15}{16}x^{-7/2}.$$

This matches the expression shown in the problem. Since the function  $f^{(4)}(x)$  is decreasing on the interval  $[1.00, 1.30]$ , the maximum absolute value occurs at the left endpoint:

$$|f^{(4)}(1.00)| = \frac{15}{16} = 0.9375.$$

Substituting into the truncation error formula:

- $b - a = 0.30$
- $h = 0.05$
- $h^4 = (0.05)^4 = 6.25 \times 10^{-8}$

The error is bounded by

$$|\text{Error}| \leq \frac{(0.30)(6.25 \times 10^{-8})}{180} \cdot \frac{15}{16} = 1.0417 \times 10^{-10}.$$

This result shows the truncation error is less than 0.00000000011, which is far smaller than the approximation 0.00000001 suggested in the book. The estimate given is conservative but valid, and the conclusion that the error is **minute** is accurate.