Numerical Integration

7.1 Introduction to Quadrature

We now approach the subject of numerical integration. The goal is to approximate the definite integral of $f(x)$ over the interval $[a, b]$ by evaluating $f(x)$ at a finite number of sample points.

Definition 7.1. Suppose that $a = x_0 < x_1 < \cdots < x_M = b$. A formula of the form

(1)
$$Q[f] = \sum_{k=0}^{M} w_k f(x_k) = w_0 f(x_0) + w_1 f(x_1) + \cdots + w_M f(x_M)$$

with the property that

(2)
$$\int_a^b f(x)\,dx = Q[f] + E[f]$$

is called a numerical integration or ***quadrature*** formula. The term $E[f]$ is called the ***truncation error*** for integration. The values $\{x_k\}_{k=0}^{M}$ are called the ***quadrature nodes*** and $\{w_k\}_{k=0}^{M}$ are called the ***weights***. ∎

Definition 7.2. The ***degree of precision*** of a quadrature formula is the positive integer n such that $E[P_i] = 0$ for all polynomials $P_i(x)$ of degree $i \leq n$, but for which $E[P_{n+1}] \neq 0$ for some polynomial $P_{n+1}(x)$ of degree $n + 1$. ▲

The form of $E[P_i]$ can be anticipated by studying what happens when $f(x)$ is a polynomial. Consider the arbitrary polynomial

$$P_i(x) = a_i x^i + a_{i-1} x^{i-1} + \cdots + a_1 x + a_0$$

of degree i. If $i \leq n$, then $P_i^{(n+1)}(x) = 0$ for all x, and $P_{n+1}^{(n+1)}(x) = (n+1)!a_{n-1}$ for all x. Thus it is not surprising that the general form for the truncation error term is

$$E[f] = K f^{(n+1)}(c), \tag{3}$$

where K is a suitably chosen constant and n is the degree of precision. The proof of this general result can be found in advanced books on numerical integration.

When we first explored finding the net signed area bounded by a curve, we developed the concept of a Riemann sum as a helpful estimation tool and a key step in the definition of the definite integral. Recall that the left, right, and middle Riemann sums of a function f on an interval $[a, b]$ are given by

$$L_n = f(x_0)\Delta x + f(x_1)\Delta x + \cdots + f(x_{n-1})\Delta x = \sum_{i=0}^{n-1} f(x_i)\Delta x, \tag{5.6.1}$$

$$R_n = f(x_1)\Delta x + f(x_2)\Delta x + \cdots + f(x_n)\Delta x = \sum_{i=1}^{n} f(x_i)\Delta x, \tag{5.6.2}$$

$$M_n = f(\overline{x}_1)\Delta x + f(\overline{x}_2)\Delta x + \cdots + f(\overline{x}_n)\Delta x = \sum_{i=1}^{n} f(\overline{x}_i)\Delta x, \tag{5.6.3}$$

where $x_0 = a$, $x_i = a + i\Delta x$, $x_n = b$, and $\Delta x = \frac{b-a}{n}$. For the middle sum, we defined $\overline{x}_i = (x_{i-1} + x_i)/2$.

A Riemann sum is a sum of (possibly signed) areas of rectangles. The value of n determines the number of rectangles, and our choice of left endpoints, right endpoints, or midpoints determines the heights of the rectangles. We can see the similarities and differences among these three options in Figure 5.6.1, where we consider the function $f(x) = \frac{1}{20}(x-4)^3 + 7$ on the interval $[1, 8]$, and use 5 rectangles for each of the Riemann sums.

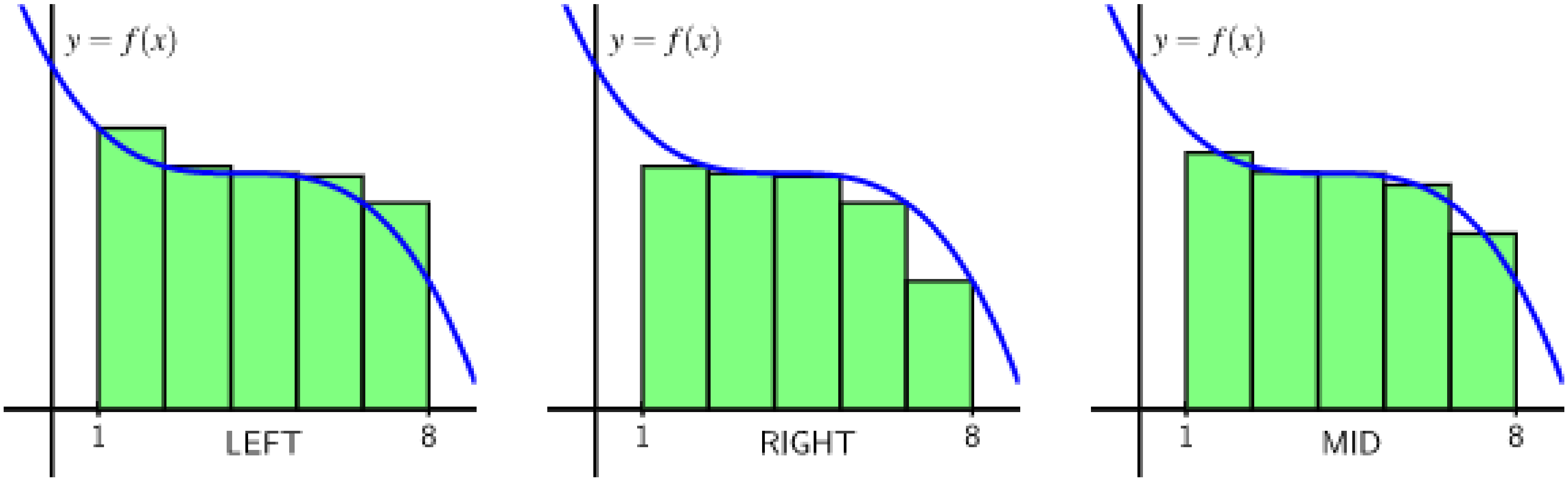


Figure 5.6.1. Left, right, and middle Riemann sums for $y = f(x)$ on $[1, 8]$ with 5 subintervals.

Midpoint Rule

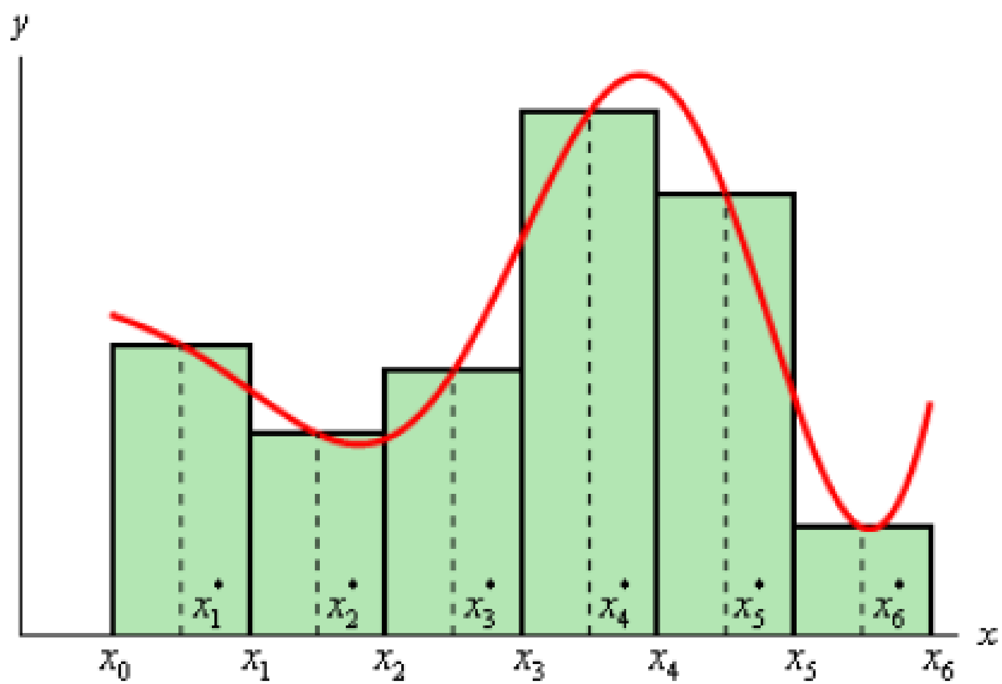
This is the rule that should be somewhat familiar to you. We will divide the interval $[a,b]$ into *n* subintervals of equal width,

$$\Delta x = \frac{b-a}{n}$$

We will denote each of the intervals as follows,

$$[x_0, x_1], [x_1, x_2], \ldots, [x_{n-1}, x_n] \quad \text{where } x_0 = a \text{ and } x_n = b$$

Then for each interval let x_i^* be the midpoint of the interval. We then sketch in rectangles for each subinterval with a height of $f(x_i^*)$. Here is a graph showing the set up using $n = 6$.



We can easily find the area for each of these rectangles and so for a general *n* we get that,

$$\int_a^b f(x)\,dx \approx \Delta x\, f(x_1^*) + \Delta x\, f(x_2^*) + \cdots + \Delta x\, f(x_n^*)$$

Or, upon factoring out a Δx we get the general Midpoint Rule.

$$\boxed{\int_a^b f(x)\,dx \approx \Delta x\left[f(x_1^*) + f(x_2^*) + \cdots + f(x_n^*)\right]}$$

Theorem 7.2 (Composite Trapezoidal Rule). Suppose that the interval $[a, b]$ is subdivided into M subintervals $[x_k, x_{k+1}]$ of width $h = (b-a)/M$ by using the equally spaced nodes $x_k = a + kh$, for $k = 0, 1, \ldots, M$. The ***composite trapezoidal rule for M subintervals*** can be expressed in any of three equivalent ways:

(1a)
$$T(f, h) = \frac{h}{2} \sum_{k=1}^{M} (f(x_{k-1}) + f(x_k))$$

or

(1b)
$$T(f, h) = \frac{h}{2}(f_0 + 2f_1 + 2f_2 + 2f_3 + \cdots + 2f_{M-2} + 2f_{M-1} + f_M)$$

or

(1c)
$$T(f, h) = \frac{h}{2}(f(a) + f(b)) + h \sum_{k=1}^{M-1} f(x_k).$$

This is an approximation to the integral of $f(x)$ over $[a, b]$, and we write

(2)
$$\int_a^b f(x)\, dx \approx T(f, h).$$

Example 7.5. Consider $f(x) = 2 + \sin(2\sqrt{x})$. Use the composite trapezoidal rule with 11 sample points to compute an approximation to the integral of $f(x)$ taken over $[1, 6]$.

To generate 11 sample points, we use $M = 10$ and $h = (6 - 1)/10 = 1/2$. Using formula (1c), the computation is

$$\begin{aligned}
T(f, \tfrac{1}{2}) &= \frac{1/2}{2}(f(1) + f(6)) \\
&\quad + \frac{1}{2}(f(\tfrac{3}{2}) + f(2) + f(\tfrac{5}{2}) + f(3) + f(\tfrac{7}{2}) + f(4) + f(\tfrac{9}{2}) + f(5) + f(\tfrac{11}{2})) \\
&= \frac{1}{4}(2.90929743 + 1.01735756) \\
&\quad + \frac{1}{2}(2.63815764 + 2.30807174 + 1.97931647 + 1.68305284 + 1.43530410 \\
&\quad + 1.24319750 + 1.10831775 + 1.02872220 + 1.00024140) \\
&= \frac{1}{4}(3.92665499) + \frac{1}{2}(14.42438165) \\
&= 0.98166375 + 7.21219083 = 8.19385457.
\end{aligned}$$

■

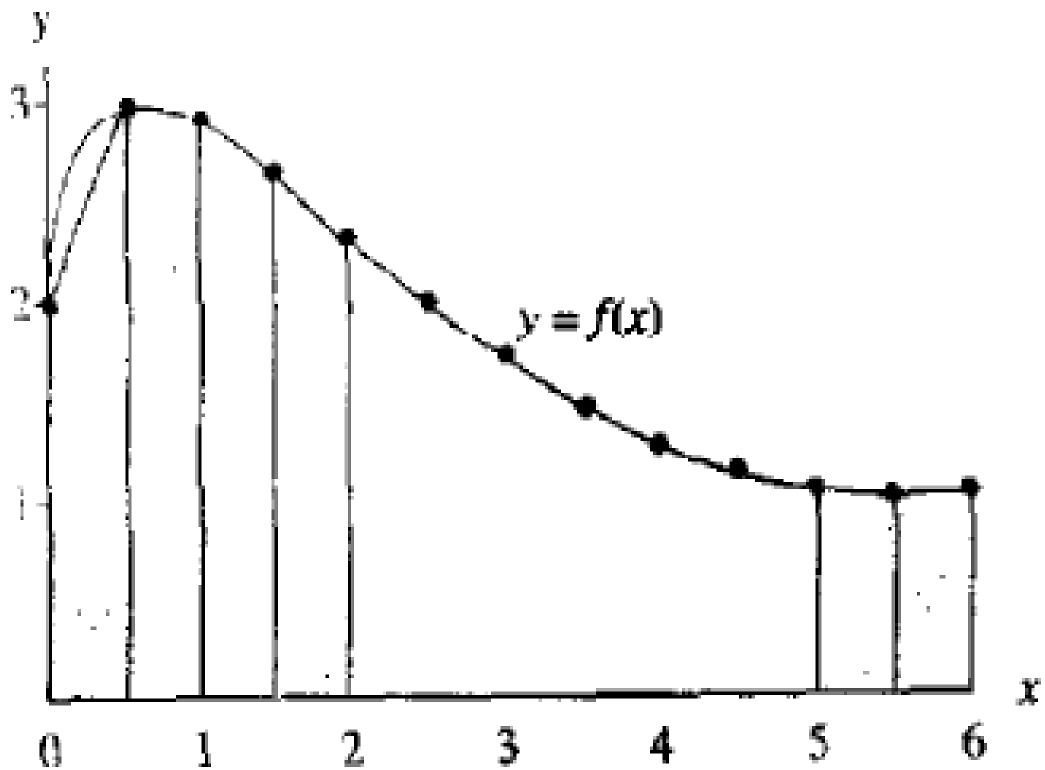


Figure 7.6 Approximating the area under the curve $y = 2 + \sin(2\sqrt{x})$ with the composite trapezoidal rule.

10.2 Simpson's Rule

The trapezoidal and Simpson's rules are special cases of the *Newton–Cote rules* which use higher degree functions for numerical integration.

Let the curve of Figure 10.4 be represented by the parabola

$$y = \alpha x^2 + \beta x + \gamma \tag{10.11}$$

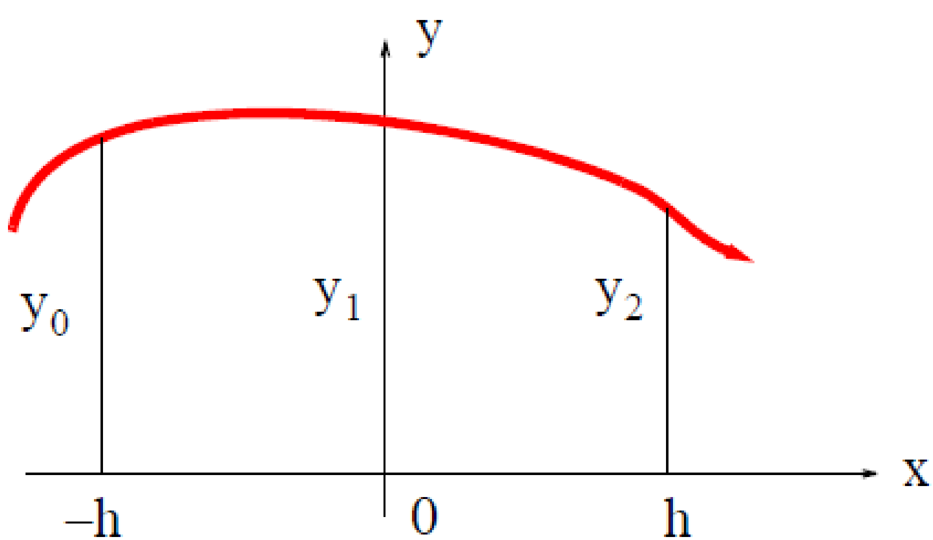


Figure 10.4. Simpson's rule of integration

The area under this curve for the interval $-h \le x \le h$ is

$$\text{Area}\Big|_{-h}^{h} = \int_{-h}^{h} (\alpha x^2 + \beta x + \gamma)dx = \frac{\alpha x^3}{3} + \frac{\beta x^2}{2} + \gamma x \Bigg|_{-h}^{h}$$

$$= \frac{\alpha h^3}{3} + \frac{\beta h^2}{2} + \gamma h - \left(-\frac{\alpha h^3}{3} + \frac{\beta h^2}{2} - \gamma h\right) = \frac{2\alpha h^3}{3} + 2\gamma h \qquad (10.12)$$

$$= \frac{1}{3}h(2\alpha h^3 + 6\gamma)$$

The curve passes through the three points $(-h, y_0)$, $(0, y_1)$, and (h, y_2). Then, by (10.11) we have:

$$\begin{aligned} y_0 &= \alpha h^2 - \beta h + \gamma && \text{(a)} \\ y_1 &= \gamma && \text{(b)} \\ y_2 &= \alpha h^2 + \beta h + \gamma && \text{(c)} \end{aligned} \qquad (10.13)$$

We can now evaluate the coefficients α, β, γ and express (10.12) in terms of h, y_0, y_1 and y_2. This is done with the following procedure.

By substitution of (b) of (10.13) into (a) and (c) and rearranging we obtain

$$\alpha h^2 - \beta h = y_0 - y_1 \tag{10.14}$$

$$\alpha h^2 + \beta h = y_2 - y_1 \tag{10.15}$$

Addition of (10.14) with (10.15) yields

$$2\alpha h^2 = y_0 - 2y_1 + y_2 \tag{10.16}$$

and by substitution into (10.12) we obtain

$$\text{Area}\Big|_{-h}^{h} = \frac{1}{3}h(2\alpha h^2 + 6\gamma) = \frac{1}{3}h[(y_0 - 2y_1 + y_2) + 6y_1] \tag{10.17}$$

or

$$\text{Area}\Big|_{-h}^{h} = \frac{1}{3}h(y_0 + 4y_1 + y_2) \tag{10.18}$$

Now, we can apply (10.18) to successive segments of any curve $y = f(x)$ in the interval $a \le x \le b$ as shown on the curve of Figure 10.5.

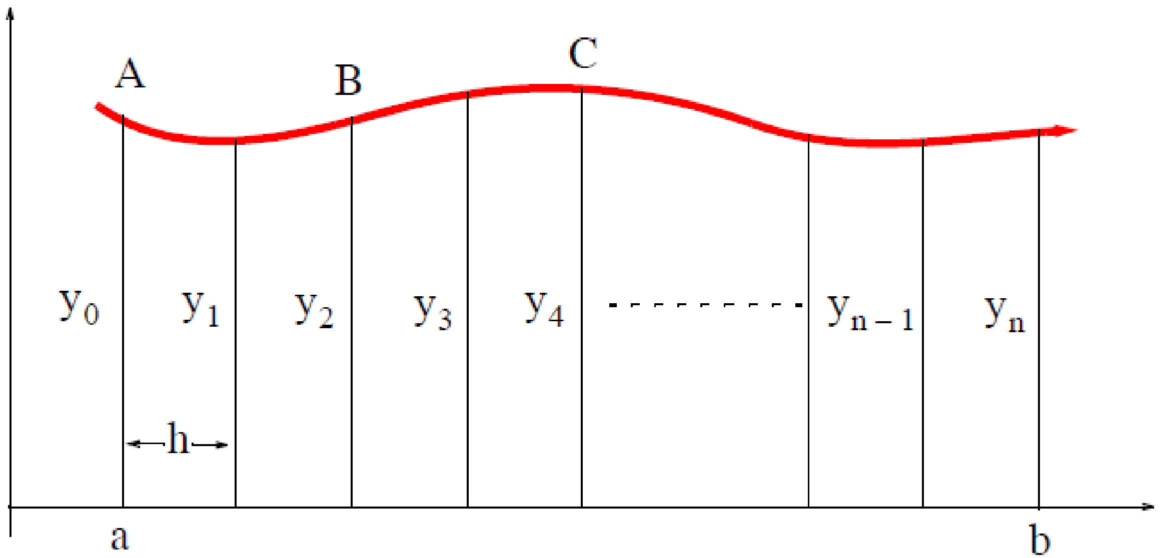


Figure 10.5. Simpson's rule of integration by successive segments

From Figure 10.5, we observe that each segment of width $2h$ of the curve can be approximated by a parabola through its ends and its midpoint. Thus, the area under segment AB is

$$\text{Area}\big|_{AB} = \frac{1}{3}h(y_0 + 4y_1 + y_2) \tag{10.19}$$

Likewise, the area under segment BC is

$$\text{Area}\big|_{BC} = \frac{1}{3}h(y_2 + 4y_3 + y_4) \tag{10.20}$$

and so on. When the areas under each segment are added, we obtain

$$\boxed{\begin{array}{c}\text{Area} = \frac{1}{3}h(y_0 + 4y_1 + 2y_2 + 4y_3 + 2y_4 + \ldots + 2y_{n-2} + 4y_{n-1} + y_n) \\ \text{Simpson's Rule of Numerical Integration}\end{array}} \tag{10.21}$$

Since each segment has width $2h$, *to apply Simpson's rule of numerical integration, the number* n *of subdivisions must be even*. This restriction does not apply to the trapezoidal rule of numerical integration. The value of h for (10.21) is found from

$$h = \frac{b-a}{n} \quad n = \text{even} \tag{10.22}$$

Theorem 7.3 (Composite Simpson Rule). Suppose that $[a, b]$ is subdivided into $2M$ subintervals $[x_k, x_{k+1}]$ of equal width $h = (b-a)/(2M)$ by using $x_k = a + kh$ for $k = 0, 1, \ldots, 2M$. The ***composite Simpson rule for 2M subintervals*** can be expressed in any of three equivalent ways:

$$S(f,h) = \frac{h}{3}\sum_{k=1}^{M}(f(x_{2k-2}) + 4f(x_{2k-1}) + f(x_{2k})) \tag{4a}$$

or

$$\begin{aligned} S(f,h) = \frac{h}{3}(&f_0 + 4f_1 + 2f_2 + 4f_3 \\ &+ \cdots + 2f_{2M-2} + 4f_{2M-1} + f_{2M}) \end{aligned} \tag{4b}$$

or

$$S(f,h) = \frac{h}{3}(f(a) + f(b)) + \frac{2h}{3}\sum_{k=1}^{M-1} f(x_{2k}) + \frac{4h}{3}\sum_{k=1}^{M} f(x_{2k-1}). \tag{4c}$$

This is an approximation to the integral of $f(x)$ over $[a, b]$, and we write

$$\int_a^b f(x)\,dx \approx S(f,h). \tag{5}$$

Example 7.6. Consider $f(x) = 2 + \sin(2\sqrt{x})$. Use the composite Simpson rule with 11 sample points to compute an approximation to the integral of $f(x)$ taken over $[1, 6]$.

To generate 11 sample points, we must use $M = 5$ and $h = (6-1)/10 = 1/2$. Using formula (4c), the computation is

$$
\begin{aligned}
S(f, \frac{1}{2}) &= \frac{1}{6}(f(1) + f(6)) + \frac{1}{3}(f(2) + f(3) + f(4) + f(5)) \\
&\quad + \frac{2}{3}(f(\tfrac{3}{2}) + f(\tfrac{5}{2}) + f(\tfrac{7}{2}) + f(\tfrac{9}{2}) + f(\tfrac{11}{2})) \\
&= \frac{1}{6}(2.90929743 + 1.01735756) \\
&\quad + \frac{1}{3}(2.30807174 + 1.68305284 + 1.24319750 + 1.02872220) \\
&\quad + \frac{2}{3}(2.63815764 + 1.97931647 + 1.43530410 + 1.10831775 + 1.00024140) \\
&= \frac{1}{6}(3.92665499) + \frac{1}{3}(6.26304429) + \frac{2}{3}(8.16133735) \\
&= 0.65444250 + 2.08768143 + 5.44089157 = 8.18301550.
\end{aligned}
$$

■

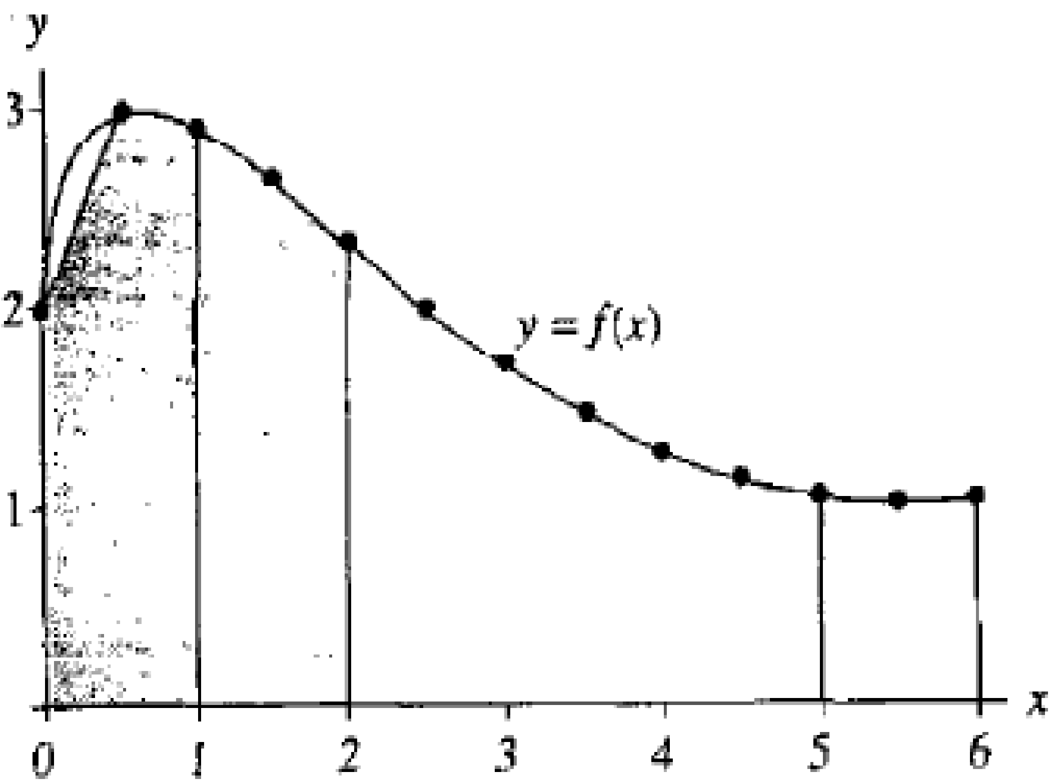


Figure 7.7 Approximating the area under the curve $y = 2 + \sin(2\sqrt{x})$ *with the composite Simpson rule.*

Corollary 7.2 (Trapezoidal Rule: Error Analysis). Suppose that $[a, b]$ is subdivided into M subintervals $[x_k, x_{k+1}]$ of width $h = (b-a)/M$. The composite trapezoidal rule

$$(7) \qquad T(f,h) = \frac{h}{2}(f(a) + f(b)) + h \sum_{k=1}^{M-1} f(x_k)$$

is an approximation to the integral

$$(8) \qquad \int_a^b f(x)\,dx = T(f,h) + E_T(f,h).$$

Furthermore, if $f \in C^2[a, b]$, there exists a value c with $a < c < b$ so that the error term $E_T(f,h)$ has the form

$$(9) \qquad E_T(f,h) = \frac{-(b-a) f^{(2)}(c) h^2}{12} = O(h^2).$$

Corollary 7.3 (Simpson's Rule: Error Analysis). Suppose that $[a, b]$ is subdivided into $2M$ subintervals $[x_k, x_{k+1}]$ of equal width $h = (b - a)/(2M)$. The composite Simpson rule

$$S(f, h) = \frac{h}{3}(f(a) + f(b)) + \frac{2h}{3}\sum_{k=1}^{M-1} f(x_{2k}) + \frac{4h}{3}\sum_{k=1}^{M} f(x_{2k-1}) \tag{14}$$

is an approximation to the integral

$$\int_a^b f(x)\,dx = S(f, h) + E_S(f, h). \tag{15}$$

Furthermore, if $f \in C^4[a, b]$, there exists a value c with $a < c < b$ so that the error term $E_S(f, h)$ has the form

$$E_S(f, h) = \frac{-(b - a) f^{(4)}(c) h^4}{180} = O(h^4). \tag{16}$$

Example 7.7. Consider $f(x) = 2 + \sin(2\sqrt{x})$. Investigate the error when the composite trapezoidal rule is used over $[1, 6]$ and the number of subintervals is 10, 20, 40, 80, and 160.

Table 7.2 shows the approximations $T(f, h)$. The antiderivative of $f(x)$ is

$$F(x) = 2x - \sqrt{x}\cos(2\sqrt{x}) + \frac{\sin(2\sqrt{x})}{2},$$

and the true value of the definite integral is

$$\int_1^6 f(x)\,dx = F(x)\Big|_{x=1}^{x=6} = 8.1834792077.$$

This value was used to compute the values $E_T(f, h) = 8.1834792077 - T(f, h)$ in Table 7.2. It is important to observe that when h is reduced by a factor of $\frac{1}{2}$ the successive errors $E_T(f, h)$ are diminished by approximately $\frac{1}{4}$. This confirms that the order is $O(h^2)$

Table 7.2 The Composite Trapezoidal Rule for $f(x) = 2 + \sin(2\sqrt{x})$ over $[1, 6]$

M	h	$T(f, h)$	$E_T(f, h) = O(h^2)$
10	0.5	8.19385457	−0.01037540
20	0.25	8.18604926	−0.00257006
40	0.125	8.18412019	−0.00064098
80	0.0625	8.18363936	−0.00016015
160	0.03125	8.18351924	−0.00004003

Example 7.8. Consider $f(x) = 2 + \sin(2\sqrt{x})$. Investigate the error when the composite Simpson rule is used over $[1, 6]$ and the number of subintervals is 10, 20, 40, 80, and 160.

Table 7.3 shows the approximations $S(f, h)$. The true value of the integral is 8.1834792077, which was used to compute the values $E_S(f, h) = 8.1834792077 - S(f, h)$ in Table 7.3. It is important to observe that when h is reduced by a factor of $\frac{1}{2}$ the successive errors $E_S(f, h)$ are diminished by approximately $\frac{1}{16}$. This confirms that the order is $O(h^4)$. ■

Table 7.3 The Composite Trapezoidal Rule for $f(x) = 2 + \sin(2\sqrt{x})$ over $[1, 6]$

M	h	$S(f, h)$	$E_S(f, h) = O(h^4)$
5	0.5	8.18301549	0.00046371
10	0.25	8.18344750	0.00003171
20	0.125	8.18347717	0.00000204
40	0.0625	8.18347908	0.00000013
80	0.03125	8.18347920	0.00000001

1. Consider integration of $f(x)$ over the fixed interval $[a, b] = [0, 1]$. Apply the various quadrature formulas (4) through (7). The step sizes are $h = 1$, $h = \frac{1}{2}$, $h = \frac{1}{3}$ and $h = \frac{1}{4}$ for the trapezoidal rule, Simpson's rule, Simpson's $\frac{3}{8}$ rule, and Boole's rule, respectively.

   (a) $f(x) = \sin(\pi x)$

   (b) $f(x) = 1 + e^{-x}\cos(4x)$

   (c) $f(x) = \sin(\sqrt{x})$

   Remark. The true values of the definite integrals are (a) $2/\pi = 0.636619772367\ldots$, (b) $(18e - \cos(4) + 4\sin(4))/(17e) = 1.007459631397\ldots$, and (c) $2(\sin(1) - \cos(1)) = 0.602337357879\ldots$. Graphs of the functions are shown in Figures 7.5(a) through (c), respectively.

2. Consider integration of $f(x)$ over the fixed interval $[a, b] = [0, 1]$. Apply the various quadrature formulas: the composite trapezoidal rule (17), the composite Simpson's rule (18), and Boole's rule (7). Use five function evaluations at equally spaced nodes. The uniform step size is $h = \frac{1}{4}$.

   (a) $f(x) = \sin(\pi x)$

   (b) $f(x) = 1 + e^{-x}\cos(4x)$

   (c) $f(x) = \sin(\sqrt{x})$

3. Consider a general interval $[a, b]$. Show that Simpson's rule produces exact results for the functions $f(x) = x^2$ and $f(x) = x^3$; that is,

   (a) $\int_a^b x^2\,dx = \frac{b^3}{3} - \frac{a^3}{3}$ (b) $\int_a^b x^3\,dx = \frac{b^4}{4} - \frac{a^4}{4}$

1. (i) Approximate each integral using the composite trapezoidal rule with $M = 10$.
(ii) Approximate each integral using the composite Simpson rule with $M = 5$.

(a) $\int_{-1}^{1} (1 + x^2)^{-1}\, dx$ (b) $\int_0^1 (2 + \sin(2\sqrt{x}))\, dx$ (c) $\int_{0.25}^{4} dx/\sqrt{x}$

(d) $\int_0^4 x^2 e^{-x}\, dx$ (e) $\int_0^2 2x \cos(x)\, dx$ (f) $\int_0^\pi \sin(2x) e^{-x}\, dx$

2. *Length of a curve.* The arc length of the curve $y = f(x)$ over the interval $a \le x \le b$ is

$$\text{length} = \int_a^b \sqrt{1 + (f'(x)^2)}\, dx.$$

(i) Approximate the arc length of each function using the composite trapezoidal rule with $M = 10$.

(ii) Approximate the arc length of each function using the composite Simpson rule with $M = 5$.

(a) $f(x) = x^3$ for $0 \le x \le 1$

(b) $f(x) = \sin(x)$ for $0 \le x \le \pi/4$

(c) $f(x) = e^{-x}$ for $0 \le x \le 1$