CHAPTER

4 Numerical Integration and Differentiation

4.1 Introduction

Many techniques are described in calculus courses for the exact evaluation of integrals, but exact techniques fail to solve many problems that arise in the physical world. For these we need approximation methods of the type we consider in this chapter. The basic techniques are discussed in Section 4.2, and refinements and special applications of these procedures are given in the next six sections.

Section 4.9 considers approximating the derivatives of functions. Methods of this type will be needed in Chapters 11 and 12 for approximating the solutions to ordinary and partial differential equations. You might wonder why there is so much more emphasis on approximating integrals than on approximating derivatives. Determining the actual derivative of a function is a constructive process that leads to straightforward rules for evaluation. Although the definition of the integral is also constructive, the principal tool for evaluating a definite integral is the Fundamental Theorem of Calculus. To apply this theorem, we must determine the antiderivative of the function we wish to evaluate. This is not generally a constructive process, and it leads to the need for accurate approximation procedures.

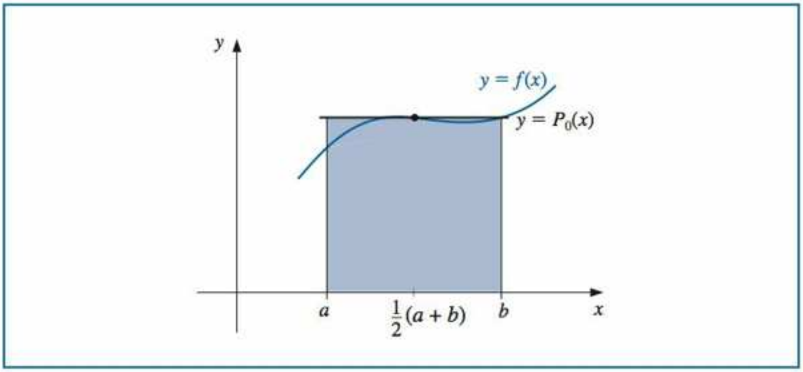
In this chapter we will also discover one of the more interesting facts in the study of numerical methods. The approximation of integrals—a task that is frequently needed—can usually be accomplished very accurately and often with little effort. The accurate approximation of derivatives—which is needed far less frequently—is a more difficult problem. We think that there is something satisfying about a subject that provides good approximation methods for problems that need them, but is less successful for problems that don't.

4.2 Basic Quadrature Rules

The basic procedure for approximating the definite integral of a function f on the interval $[a, b]$ is to determine an interpolating polynomial that approximates f and then integrate this polynomial. In this section we determine approximations that arise when some basic polynomials are used for the approximations and determine error bounds for these approximations.

The approximations we consider use interpolating polynomials at equally spaced points in the interval $[a, b]$. The first of these is the *Midpoint rule*, which uses the midpoint of $[a, b]$, $\frac{1}{2}(a+b)$, as its only interpolation point. The Midpoint rule approximation is easy to generate geometrically, as shown in Figure 4.1, but to establish the pattern for the higher-order methods and to determine an error formula for the technique, we will use a basic tool for these derivations, the Newton interpolatory divided-difference formula which we discussed on page 76.

Figure 4.1



Suppose that $f \in C^{n+1}[a,b]$, where $[a,b]$ is an interval that contains all the nodes $x_0, x_1, \ldots, x_n$. The Newton interpolatory divided-difference formula states that the interpolating polynomial for the function f using the nodes $x_0, x_1, \ldots, x_n$ can be expressed in the form

$$\begin{aligned}P_{0,1,\ldots,n}(x) = f[x_0] + f[x_0,x_1](x-x_0) + f[x_0,x_1,x_2](x-x_0)(x-x_1) + \cdots \\ + f[x_0,x_1,\ldots,x_n](x-x_0)(x-x_1)\cdots(x-x_{n-1}).\end{aligned}$$

Since this is equivalent to the nth Lagrange polynomial, the error formula has the form

$$f(x) - P_{0,1,\ldots,n}(x) = \frac{f^{(n+1)}(\xi(x))}{(n+1)!}(x-x_0)(x-x_1)\cdots(x-x_n),$$

where $\xi(x)$ is a number, depending on x, that lies in the smallest interval that contains all of $x, x_0, x_1, \ldots, x_n$.

To derive the Midpoint rule we could use the constant interpolating polynomial with $x_0 = \frac{1}{2}(a+b)$ to produce

$$\int_a^b f(x)dx \approx \int_a^b f[x_0]dx = f[x_0](b-a) = f\left(\frac{a+b}{2}\right)(b-a).$$

But we could also use a linear interpolating polynomial with this value of x_0 and an arbitrary value of x_1. This is because the integral of the second term in the Newton interpolatory divided-difference formula is zero for our choice of x_0, independent of the value of x_1, and as such does not contribute to the approximation:

$$\begin{aligned}\int_a^b f[x_0,x_1](x-x_0)dx &= \frac{f[x_0,x_1]}{2}(x-x_0)^2\Big]_a^b \\ &= \frac{f[x_0,x_1]}{2}\left(x-\frac{a+b}{2}\right)^2\Big]_a^b \\ &= \frac{f[x_0,x_1]}{2}\left[\left(b-\frac{a+b}{2}\right)^2 - \left(a-\frac{a+b}{2}\right)^2\right] \\ &= \frac{f[x_0,x_1]}{2}\left[\left(\frac{b-a}{2}\right)^2 - \left(\frac{a-b}{2}\right)^2\right] = 0.\end{aligned}$$

We would like to derive approximation methods that have high powers of $b-a$ in the error term. In general, the higher the degree of the approximation, the higher the power of $b-a$ in the error term, so we will integrate the error for the linear interpolation polynomial instead of the constant polynomial to determine an error formula for the Midpoint rule.

Suppose that the arbitrary x_1 was chosen to be the same value as x_0. (In fact, this is the only value that we *cannot* have for x_1, but we will ignore this problem for the moment.) Then the integral of the error formula for the interpolating polynomial $P_{0,1}(x)$ has the form

$$\int_a^b \frac{(x-x_0)(x-x_1)}{2} f''(\xi(x))dx = \int_a^b \frac{(x-x_0)^2}{2} f''(\xi(x))dx,$$

where, for each x, the number $\xi(x)$ lies in the interval (a, b).

The term $(x-x_0)^2$ does not change sign on the interval (a, b), so the Mean Value Theorem for Integrals (see page 8) implies that a number ξ, independent of x, exists in (a, b) with

$$\begin{aligned}\int_a^b \frac{(x-x_0)^2}{2} f''(\xi(x))dx &= f''(\xi)\int_a^b \frac{(x-x_0)^2}{2}dx = \frac{f''(\xi)}{6}(x-x_0)^3\Big]_a^b\\ &= \frac{f''(\xi)}{6}\left[\left(b-\frac{b+a}{2}\right)^3-\left(a-\frac{b+a}{2}\right)^3\right]\\ &= \frac{f''(\xi)}{6}\frac{(b-a)^3}{4} = \frac{f''(\xi)}{24}(b-a)^3.\end{aligned}$$

As a consequence, the Midpoint rule with its error formula has the following form:

Midpoint Rule

If $f \in C^2[a, b]$, then a number ξ in (a, b) exists with

$$\int_a^b f(x)dx = (b-a)f\left(\frac{a+b}{2}\right) + \frac{f''(\xi)}{24}(b-a)^3.$$

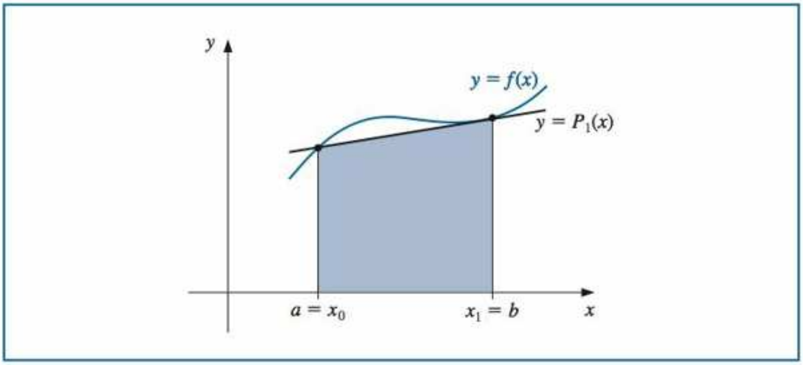
The invalid assumption, $x_1 = x_0$, that leads to this result can be avoided by taking x_1 close, but not equal, to x_0 and using limits to show that the error formula is still valid.

The Trapezoidal Rule

The Midpoint rule uses a constant interpolating polynomial disguised as a linear interpolating polynomial. The next method we consider uses a true linear interpolating polynomial, one with the distinct nodes $x_0 = a$ and $x_1 = b$. This approximation is also easy to generate geometrically, as shown in Figure 4.2 on the following page, and is aptly called the *Trapezoidal*, or *Trapezium*, rule. If we integrate the linear interpolating polynomial with $x_0 = a$ and $x_1 = b$, we also produce this formula:

$$\begin{aligned}\int_a^b f[x_0] + f[x_0, x_1](x-x_0)dx &= \left[f[a]x + f[a,b]\frac{(x-a)^2}{2}\right]_a^b\\ &= f(a)(b-a) + \frac{f(b)-f(a)}{b-a}\left[\frac{(b-a)^2}{2} - \frac{(a-a)^2}{2}\right]\\ &= \frac{f(a)+f(b)}{2}(b-a).\end{aligned}$$

Figure 4.2



The error for the Trapezoidal rule follows from integrating the error term for $P_{0,1}(x)$ when $x_0 = a$ and $x_1 = b$. Since $(x - x_0)(x - x_1) = (x - a)(x - b)$ is always negative in the interval (a, b), we can again apply the Mean Value Theorem for Integrals. In this case it implies that a number ξ in (a, b) exists with

$$\begin{aligned}\int_a^b \frac{(x-a)(x-b)}{2} f''(\xi(x))dx &= \frac{f''(\xi)}{2}\int_a^b (x-a)[(x-a)-(b-a)]dx\\ &= \frac{f''(\xi)}{2}\left[\frac{(x-a)^3}{3} - \frac{(x-a)^2}{2}(b-a)\right]_a^b\\ &= \frac{f''(\xi)}{2}\left[\frac{(b-a)^3}{3} - \frac{(b-a)^2}{2}(b-a)\right]\\ &= -\frac{f''(\xi)}{12}(b-a)^3.\end{aligned}$$

This gives the Trapezoidal rule with its error formula.

Trapezoidal Rule

If $f \in C^2[a, b]$, then a number ξ in (a, b) exists with

$$\int_a^b f(x)dx = \frac{f(a)+f(b)}{2}(b-a) - \frac{f''(\xi)}{12}(b-a)^3.$$

When we use the term *trapezoid* we mean a four-sided figure that has at least two of its sides parallel. The European term for this figure is *trapezium*. To further confuse the issue, the European word trapezoidal refers to a four-sided figure with no sides equal, and the American word for this type of figure is trapezium.

We cannot improve on the power of $b - a$ in the error formula for the Trapezoidal rule, as we did in the case of the Midpoint rule, because the integral of the next higher term in the Newton interpolatory divided-difference formula is

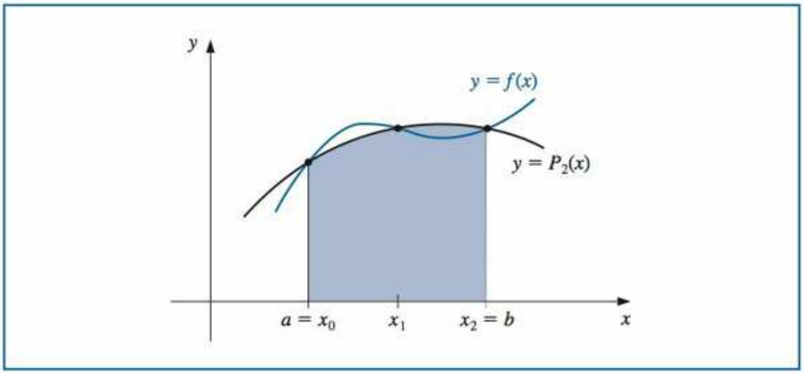
$$\int_a^b f[x_0, x_1, x_2](x - x_0)(x - x_1)dx = f[x_0, x_1, x_2]\int_a^b (x-a)(x-b)dx.$$

Since $(x-a)(x-b) < 0$ for all x in (a, b), this term will not be zero unless $f[x_0, x_1, x_2] = 0$. As a consequence, the error formulas for the Midpoint and the Trapezoidal rules both involve $(b - a)^3$, even though they are derived from interpolation formulas with error formulas that involve $b - a$ and $(b - a)^2$, respectively.

Simpson's Rule

Next we consider an integration formula based on approximating the function f by a quadratic polynomial that agrees with f at the equally spaced points $x_0 = a, x_1 = (a+b)/2$, and $x_2 = b$. This formula is not easy to generate geometrically, although the approximation is illustrated in Figure 4.3.

Figure 4.3



To derive the formula, we integrate $P_{0,1,2}(x)$.

$$
\begin{aligned}
&\int_a^b P_{0,1,2}(x)dx \\
&\quad = \int_a^b \left\{ f(a) + f\left[a, \frac{a+b}{2}\right](x-a) + f\left[a, \frac{a+b}{2}, b\right](x-a)\left(x - \frac{a+b}{2}\right)\right\} dx \\
&\quad = \left[f(a)x + f\left[a, \frac{a+b}{2}\right] \frac{(x-a)^2}{2}\right]_a^b \\
&\qquad + f\left[a, \frac{a+b}{2}, b\right] \int_a^b (x-a)\left[(x-a) + \left(a - \frac{a+b}{2}\right)\right] dx \\
&\quad = f(a)(b-a) + \frac{f(\frac{a+b}{2}) - f(a)}{\frac{a+b}{2} - a} \frac{(b-a)^2}{2} \\
&\qquad + \frac{f[\frac{a+b}{2}, b] - f[a, \frac{a+b}{2}]}{b-a} \left[\frac{(x-a)^3}{3} + \frac{(x-a)^2}{2}\left(\frac{a-b}{2}\right)\right]_a^b \\
&\quad = (b-a)\left[f(a) + f\left(\frac{a+b}{2}\right) - f(a)\right] \\
&\qquad + \left(\frac{1}{b-a}\right)\left[\frac{f(b) - f(\frac{a+b}{2})}{\frac{b-a}{2}} - \frac{f(\frac{a+b}{2}) - f(a)}{\frac{b-a}{2}}\right]\left[\frac{(b-a)^3}{3} - \frac{(b-a)^3}{4}\right] \\
&\quad = (b-a)f\left(\frac{a+b}{2}\right) + \frac{2}{(b-a)^2}\left[f(b) - 2f\left(\frac{a+b}{2}\right) + f(a)\right]\frac{(b-a)^3}{12}.
\end{aligned}
$$

Thomas Simpson (1710–1761) was a self-taught mathematician who supported himself as a weaver during his early years. His primary interest was probability theory, although in 1750 he published a two-volume calculus book entitled *The Doctrine and Application of Fluxions*.

Simplifying this equation gives the approximation method known as Simpson's rule:

$$\int_a^b f(x)dx \approx \frac{(b-a)}{6}\left[f(a) + 4f\left(\frac{a+b}{2}\right) + f(b)\right].$$

An error formula for Simpson's rule involving $(b-a)^4$ can be derived by using the error formula for the quadratic interpolating polynomial $P_{0,1,2}(x)$. However, similar to the case of the Midpoint rule, the integral of the next term in the Newton interpolatory divided-difference formula is zero. This implies that the error formula for the cubic interpolating polynomial $P_{0,1,2,3}(x)$ can be used to produce an error formula that involves $(b-a)^5$. When simplified, Simpson's rule with this error formula is as follows:

Simpson's Rule

If $f \in C^4[a, b]$, then a number ξ in (a, b) exists with

$$\int_a^b f(x)dx = \frac{(b-a)}{6}\left[f(a) + 4f\left(\frac{a+b}{2}\right) + f(b)\right] - \frac{f^{(4)}(\xi)}{2880}(b-a)^5.$$

This higher power of $b-a$ in the error term makes Simpson's rule significantly superior to the Midpoint and Trapezoidal rules in almost all situations, provided that $b-a$ is small. This is illustrated in the following example.

Example 1 Compare the Midpoint, Trapezoidal, and Simpson's rules approximations to $\int_0^2 f(x)dx$ when $f(x)$ is

(a) x^2 **(b)** x^4 **(c)** $(x+1)^{-1}$
(d) $\sqrt{1+x^2}$ **(e)** $\sin x$ **(f)** e^x

Solution On [0, 2] the Midpoint, Trapezoidal, and Simpson's rules have the forms

$$\text{Midpoint: } \int_0^2 f(x)dx \approx 2f(1), \quad \text{Trapezoidal: } \int_0^2 f(x)dx \approx f(0) + f(2),$$

and

$$\text{Simpson's: } \int_0^2 f(x)dx \approx \frac{1}{3}[f(0) + 4f(1) + f(2)].$$

When $f(x) = x^2$, they give

$$\text{Midpoint: } \int_0^2 f(x)dx \approx 2 \cdot 1 = 2, \quad \text{Trapezoidal: } \int_0^2 f(x)dx \approx 0^2 + 2^2 = 4,$$

and

$$\text{Simpson's: } \int_0^2 f(x)dx \approx \frac{1}{3}[f(0) + 4f(1) + f(2)] = \frac{1}{3}(0^2 + 4 \cdot 1^2 + 2^2) = \frac{8}{3}.$$

The approximation from Simpson's rule is exact because its truncation error involves $f^{(4)}$, which is identically 0 when $f(x) = x^2$.

The results to three places for the functions are summarized in Table 4.1. Notice that, in each instance, Simpson's rule is significantly superior. ▪

Table 4.1

	(a)	(b)	(c)	(d)	(e)	(f)
$f(x)$	x^2	x^4	$(x+1)^{-1}$	$\sqrt{1+x^2}$	$\sin x$	e^x
Exact value	2.667	6.400	1.099	2.958	1.416	6.389
Midpoint	2.000	2.000	1.000	2.818	1.682	5.436
Trapezoidal	4.000	16.000	1.333	3.326	0.909	8.389
Simpson's	2.667	6.667	1.111	2.964	1.425	6.421

To demonstrate the error terms for the Midpoint, Trapezoidal, and Simpson's methods, we will find bounds for the errors in approximating $\int_0^2 \sqrt{1+x^2}dx$. With $f(x) = (1+x^2)^{1/2}$, we have

$$f'(x) = \frac{x}{(1+x^2)^{1/2}}, \quad f''(x) = \frac{1}{(1+x^2)^{3/2}}, \quad \text{and} \quad f'''(x) = \frac{-3x}{(1+x^2)^{5/2}}.$$

To bound the error of the Midpoint method, we need to determine $\max_{0\le x\le 2}|f''(x)|$. This maximum will occur at either the maximum or the minimum value of f'' on $[0, 2]$. Maximum and minimum values for f'' on $[0, 2]$ can occur only when $x = 0, x = 2$, or when $f'''(x) = 0$. Since $f'''(x) = 0$ only when $x = 0$, we have

$$\max_{0\le x\le 2}|f''(x)| = \max\{|f''(0)|, |f''(2)|\} = \max\left\{1, 5^{-3/2}\right\} = 1.$$

So a bound for the error in the Midpoint method is

$$\left|\frac{f''(\xi)}{24}(b-a)^3\right| \le \frac{1}{24}(2-0)^3 = \frac{1}{3} = 0.\bar{3}.$$

The actual error is within this bound, since $|2.958 - 2.818| = 0.14$. For the Trapezoidal method, we have the error bound

$$\left|-\frac{f''(\xi)}{12}(b-a)^3\right| \le \frac{1}{12}(2-0)^3 = \frac{2}{3} = 0.\bar{6},$$

and the actual error is $|2.958 - 3.326| = 0.368$. We need more derivatives for Simpson's rule:

$$f^{(4)}(x) = \frac{12x^2-3}{(1+x^2)^{7/2}} \quad \text{and} \quad f^{(5)}(x) = \frac{45x-60x^3}{(1+x^2)^{9/2}}.$$

Since $f^{(5)}(x) = 0$ implies

$$0 = 45x - 60x^3 = 15x(3-4x^2),$$

$f^{(4)}(x)$ has critical points $0, \pm\sqrt{3}/2$. Evaluating the fourth derivative at the critical points and endpoints we have

$$|f^{(4)}(\xi)| \le \max_{0\le x\le 2}|f^{(4)}(x)| = \max\{|f^{(4)}(0)|, |f^{(4)}(\sqrt{3}/2)|, |f^{(4)}(2)|\}$$

$$= \max\left\{|-3|, \frac{768\sqrt{7}}{2401}, \frac{9\sqrt{5}}{125}\right\} = 3.$$

The error for Simpson's rule is consequently bounded by

$$\left| -\frac{f^{(4)}(\xi)}{2880}(b-a)^5 \right| \le \frac{3}{2880}(2-0)^5 = \frac{96}{2880} = 0.0\overline{3},$$

and the actual error is $|2.958 - 2.964| = 0.006$.

The error formulas all contain $b - a$ to a power, so they are most effective when the interval $[a, b]$ is small, so that $b - a$ is much smaller than one. There are formulas that can be used to improve the accuracy when integrating over large intervals, some of which are considered in the exercises. However, a better solution to the problem is considered in the next section.

EXERCISE SET 4.2

1. Use the Midpoint rule to approximate the following integrals.

   a. $\int_{0.5}^{1} x^4 dx$ b. $\int_0^{0.5} \frac{2}{x-4} dx$

   c. $\int_1^{1.5} x^2 \ln x \, dx$ d. $\int_0^1 x^2 e^{-x} dx$

   e. $\int_1^{1.6} \frac{2x}{x^2-4} dx$ f. $\int_0^{0.35} \frac{2}{x^2-4} dx$

   g. $\int_0^{\pi/4} x \sin x \, dx$ h. $\int_0^{\pi/4} e^{3x} \sin 2x \, dx$

2. Use the error formula to find a bound for the error in Exercise 1, and compare the bound to the actual error.
3. Repeat Exercise 1 using the Trapezoidal rule.
4. Repeat Exercise 2 using the Trapezoidal rule and the results of Exercise 3.
5. Repeat Exercise 1 using Simpson's rule.
6. Repeat Exercise 2 using Simpson's rule and the results of Exercise 5.

   Other quadrature formulas with error terms are given by

   (i) $\int_a^b f(x)dx = \frac{3h}{8}[f(a) + 3f(a+h) + 3f(a+2h) + f(b)] - \frac{3h^5}{80} f^{(4)}(\xi)$, where $h = \frac{b-a}{3}$;

   (ii) $\int_a^b f(x)dx = \frac{3h}{2}[f(a+h) + f(a+2h)] + \frac{3h^3}{4} f''(\xi)$, where $h = \frac{b-a}{3}$.

7. Repeat Exercises 1 using (a) Formula (i) and (b) Formula (ii).
8. Repeat Exercises 2 using (a) Formula (i) and (b) Formula (ii).
9. The Trapezoidal rule applied to $\int_0^2 f(x)dx$ gives the value 4, and Simpson's rule gives the value 2. What is $f(1)$?
10. The Trapezoidal rule applied to $\int_0^2 f(x)dx$ gives the value 5, and the Midpoint rule gives the value 4. What value does Simpson's rule give?
11. Find the constants c_0, c_1, and x_1 so that the quadrature formula

   $$\int_0^1 f(x)dx = c_0 f(0) + c_1 f(x_1)$$

   gives exact results for all polynomials of degree at most 2.
12. Find the constants x_0, x_1, and c_1 so that the quadrature formula

   $$\int_0^1 f(x)dx = \frac{1}{2} f(x_0) + c_1 f(x_1)$$

   gives exact results for all polynomials of degree at most 3.

13. Given the function f at the following values:

x	1.8	2.0	2.2	2.4	2.6
$f(x)$	3.12014	4.42569	6.04241	8.03014	10.46675

a. Approximate $\int_{1.8}^{2.6} f(x)dx$ using each of the following.

(i) the Midpoint rule **(ii)** the Trapezoidal rule **(iii)** Simpson's rule

b. Suppose the data have round-off errors given by the following table:

x	1.8	2.0	2.2	2.4	2.6
Error in $f(x)$	2×10^{-6}	-2×10^{-6}	-0.9×10^{-6}	-0.9×10^{-6}	2×10^{-6}

Calculate the errors due to round-off in each of the approximation methods.

4.3 Composite Quadrature Rules

The basic notions underlying numerical integration were derived in the previous section, but the techniques given there are not satisfactory for most problems. We saw an example of this at the end of that section, where the approximations were poor for integrals of functions on the interval $[0, 2]$. To see why this occurs, let us consider Simpson's method, generally the most accurate of these techniques. Assuming that $f \in C^4[a, b]$, Simpson's method with its error formula is given by

Piecewise approximation is often effective. Recall that this was used for spline interpolation.

$$\int_a^b f(x)dx = \frac{b-a}{6}\left[f(a) + 4f\left(\frac{a+b}{2}\right) + f(b)\right] - \frac{(b-a)^5}{2880}f^{(4)}(\xi)$$
$$= \frac{h}{3}[f(a) + 4f(a+h) + f(b)] - \frac{h^5}{90}f^{(4)}(\xi)$$

where $h = (b-a)/2$ and ξ lies somewhere in the interval (a, b). Since $f \in C^4[a, b]$ implies that $f^{(4)}$ is bounded on $[a, b]$, there exists a constant M such that $|f^{(4)}(x)| \le M$ for all x in $[a, b]$. As a consequence,

$$\left|\frac{h}{3}[f(a) + 4f(a+h) + f(b)] - \int_a^b f(x)dx\right| = \left|\frac{h^5}{90}f^{(4)}(\xi)\right| \le \frac{M}{90}h^5.$$

The error term in this formula involves M, a bound for the fourth derivative of f, and h^5, so we can expect the error to be small provided that

- the fourth derivative of f is not erratic, and
- the value of $h = b - a$ is small.

The first assumption we will need to live with, but the second might be quite unreasonable. There is no reason, in general, to expect that the interval $[a, b]$ over which the integration is performed is small, and if it is not, the h^5 portion in the error term will likely dominate the calculations.

We circumvent the problem involving a large interval of integration by subdividing the interval $[a, b]$ into a collection of intervals that are sufficiently small so that the error over each is kept under control.