LECTURE I: PRIMITIVE AND COMPOUND INTEGRATION

1.1. Approximate Integration Over a Finite Interval

The basic problem in numerical integration is to compute an approximate solution to a definite integral

$$I(f)=\int_a^b f(x)dx, \tag{1.1}$$

to a given degree of accuracy. If $f(x)$ is a smooth function integrated over the domain, and the domain of integration is bounded, there are many methods for approximating the integral to the desired precision.

By a primitive rule of approximate numerical integration for Eq. (1.1) we mean either a Riemann sum of the form

$$R_n(f)=R_n=h\sum_{k=1}^{n} f(a+kh)=h\sum_{k=0}^{n-1} f(a+kh),\ h=\frac{b-a}{n}, \tag{1.2}$$

this is called rectangles rule,

$$M_n(f)=M_n=h\sum_{k=0}^{n-1} f\left(a+\left(k+\frac{1}{2}\right)h\right),\ h=\frac{b-a}{n}, \tag{1.3}$$

the midpoint rule,

$$T_n(f)=T_n=h\sum_{k=0}^{n}{}'' f(a+kh)=h\left[\frac{f(a)}{2}+f(a+h)+\ldots+f(a+(n-1)h)+\frac{f(b)}{2}\right], \tag{1.4}$$

the trapezoidal rule.

Trapezoidal rule for non-equidistant point can be written in the form

$$T(f;\pi_n)=\sum_{i=1}^{n}\frac{x_1-x_{i-1}}{2}\left(f(x_{i-1})+f(x_i)\right),$$

where the partition π_n is given by $\pi_n: a=x_0<x_1<\ldots<x_n=b.$

If $x_i=a+ih,\ h=\dfrac{b-a}{n}$ then $T(f;\pi_n)$ reduces to $T_n(f)$. A similar situation holds for the midpoint rule and rectangle rule.

The Riemann sums are simple averages of the function. Note also that the trapezoidal rule is the average of the "right-hand" and the "left-hand" Riemann sums for functions periodic over $[a,b]$. The convergence to the integral of primitive sums is not rapid (exceptions must be made in the case of certain periodic functions) and is governed by the following estimates of error.

Definition 1.1: Let $f(x)$ be continuous in $[a,b]$. Then the *modulus of continuity* $w(\delta)$ of $f(x)$ is defined by

$$w(\delta)=\max_{|x_1-x_2|\le\delta}\left|f(x_1)-f(x_2)\right|,\ \ a\le x_1,x_2\le b.\,, \tag{1.5}$$

In other words, the inequality $|x_1-x_2|\le\delta$ implies that

$$\left|f(x_1)-f(x_2)\right|\le w(\delta).$$

As the interval δ becomes smaller, the variation of $f(x)$ becomes smaller and smaller so that

$$\lim_{\delta\to 0} w(\delta)=0.$$

THEOREM 1.2: Let $f(x)$ be continuous in $[a,b]$ and $x_k=a+kh,\ k=0,1,\ldots,n,\ \ h=\dfrac{b-a}{n}$, as well as $\bar{x}_k=x_k+\dfrac{h}{2},\ k=0,1,\ldots,n-1$. Then

$$\left|I(f)-R_n(f)\right|=\left|\int_a^b f(x)dx-h\sum_{k=1}^{n} f(x_k)\right|\le(b-a)w\left(\frac{b-a}{n}\right),$$

for rectangle rule

$$\left|I(f)-M_n(f)\right|=\left|\int_a^b f(x)dx-h\sum_{k=0}^{n-1} f(\bar{x}_k)\right|\le(b-a)w\left(\frac{b-a}{2n}\right),$$

for midpoint rule

$$\left|I(f)-T_n(f)\right|=\left|\int_a^b f(x)dx-h\sum_{k=0}^{n}{}'' f(x_k)\right|\le(b-a)w\left(\frac{b-a}{2n}\right),$$

for trapezoidal rule.

Proof: For rectangle rule,

$$I(f)-I_n(f)=\int_a^b f(x)dx-h\sum_{k=1}^{n} f(x_k)=\sum_{k=1}^{n}\int_{x_{k-1}}^{x_k}\left(f(x)-f(x_k)\right)dx.$$

Since $\left|f(x)-f(x_k)\right|\le w(h)$ for $x_{k-1}\le x\le x_k$ we have

$$\left|\int_{x_{k-1}}^{x_k}\left(f(x)-f(x_k)\right)dx\right|\le hw(h).$$

Hence, by adding *n* of these inequalities, we obtain

$$I(f)-I_n(f)=\left|\int_a^b f(x)dx-h\sum_{k=1}^{n} f(x_k)\right|\le nhw(h)=(b-a)w\left(\frac{b-a}{n}\right).$$

For midpoint rule,

$$I(f)-M_n(f)=\int_a^b f(x)dx-h\sum_{k=0}^{n-1} f(\bar{x}_k)=\sum_{k=0}^{n-1}\int_{x_k}^{x_{k+1}}\left(f(x)-f(\bar{x}_k)\right)dx\,,$$

Since $\left|f(x)-f(\bar{x}_k)\right|\le w\left(\dfrac{h}{2}\right)$ for $x_{k-1}\le x\le x_k$ we have

$$\left|\int_{x_{k-1}}^{x_k}\left(f(x)-f(x_k)\right)dx\right|\le hw\left(\frac{h}{2}\right).$$

Hence, by adding n of these inequalities, we obtain

$$I(f)-M_n(f)=\left|\int_a^b f(x)dx-h\sum_{k=1}^{n}f(\bar{x}_k)\right|\le nhw\left(\frac{h}{2}\right)=(b-a)w\left(\frac{b-a}{n}\right).$$

For trapezoidal rule we construct quadrature rule in different form

$$\begin{aligned}I(f)-T_n(f)&=\sum_{k=0}^{n-1}\int_{x_k}^{t_{k+1}}\left[f(x)-\frac{x-x_k}{h}f(x_{k+1})-\frac{x_{k+1}-x}{h}f(x_k)\right]dx\\&=\sum_{k=0}^{n-1}\int_{x_k}^{t_{k+1}}\frac{1}{h}\left[(x-x_k)\left(f(x)-f(x_{k+1})\right)+(x_{k+1}-x)\left(f(x)-f(x_k)\right)\right]dx\end{aligned},$$

As we know that $\left|f(x)-f(x_k)\right|\le w(h)$ and $\int_{x_k}^{x_{k+1}}\frac{x-x_k}{h}dx=\frac{h}{2}$, then

$$\begin{aligned}\left|I(f)-T_n(f)\right|&\le\sum_{k=0}^{n-1}\int_{x_k}^{t_{k+1}}\frac{1}{h}\left|(x-x_k)\left(f(x)-f(x_{k+1})\right)+(x_{k+1}-x)\left(f(x)-f(x_k)\right)\right|dx\\&\le\sum_{k=0}^{n-1}\left[\frac{h}{2}w(h)+\frac{h}{2}w(h)\right]=(b-a)w\left(\frac{b-a}{n}\right)\end{aligned}$$

These estimate tell us that for continuous functions $f(x)$ the sum (1.2) - (1.4) approach the integral $\int_a^b f(x)dx$ with rapidity of $w\left[(b-a)/n\right]$ approaching 0.

More precise error estimates can be given when restrictions are made on the smoothness of the integrand function. We have, for example, the following.

Theorem 1.3: Let

$$E_n(f)=E_n=\int_a^b f(x)dx-h\sum_{k=1}^{n}f(a+kh),\tag{1.6}$$

If $f'(x)$ exists and is bounded in $[a,b]$, then

$$\tfrac{1}{2}h^2\sum_{k=1}^{n}m_k\le E_n\le\tfrac{1}{2}h^2\sum_{k=1}^{n}M_k,\tag{1.7}$$

where $m_k=\inf_{x_{k-1}\le x\le x_k}\left|f'(x)\right|$, $M_k=\sup_{x_{k-1}\le x\le x_k}\left|f'(x)\right|$. If it is further assumed that $f'(x)$ is integrable over $[a,b]$, then we have

$$\lim_{n\to\infty}nE_n=\frac{b-a}{2}\left[f(b)-f(a)\right].,\tag{1.8}$$

For such functions, therefore, if we assume that $f(a)\ne f(b)$ the error in the Riemann sum goes to 0 precisely as 1/n. The trapezoidal and midpoint rules are exact for linear functions and converge at lest as fast as n^{-2}, if we assume that the integrand function has a continuous second derivative.

Theorem 1.4: Let $f(x)\in C^2[a,b]$. Then error term of midpoint and trapezoidal rules are

$$I(f)-M_n(f)=\int_a^b f(x)dx-h\sum_{k=0}^{n-1} f(\bar{x}_k)=\frac{(b-a)^3}{24n^2}f''(\xi),\ \ a<\xi<b., \tag{1.9}$$

$$I(f)-T_n(f)=\int_a^b f(x)dx-h\left[\frac{f(a)}{2}+\sum_{k=1}^{n-1} f(x_k)+\frac{f(b)}{2}\right]=-\frac{(b-a)^3}{12n^2}f''(\xi),\qquad a<\xi<b,, \tag{1.10}$$

Corollary 1.5: If $f(x)\in C^2[a,b]$ and $f''(x)\geq 0$ on $[a,b]$, then

$$M_n(f)\leq\int_a^b f(x)dx\leq T_n(f).$$

This is known as the "*bracketing*" property.

The Corollary 1.5 also holds if $f(x)$ is convex function on $[a,b]$. In addition, it holds for $M_n(f)$ and $T_m(f)$ with any pair of integers *m,n*. From this it follows that $T_{2n}=\frac{1}{2}(T_n+M_n)$, and

$$T_n(f)-\int_a^b f(x)dx\geq\int_a^b f(x)dx-M_n(f).$$

It shows that for convex functions, the midpoint rule is better than the corresponding trapezoidal rule.

Proof of the Theorem 1.4. Since $f(x)\in C^2[a,b]$ we can use Taylor formulas

$$f(x)=f(x_k)+f'(x_k)(x-x_k)+\frac{f''(\xi)}{2}(x-x_k)^2,\ x\in[x_k,x_{k+1}].$$

For the **midpoint rule**

$$f(x)-f(\bar{x}_k)=f'(\bar{x}_k)(x-\bar{x}_k)+\frac{f''(\xi)}{2}(x-\bar{x}_k)^2,\ \bar{x}_k=\frac{x_k+x_{k+1}}{2},\ x\in[x_k,x_{k+1}].$$

We have

$$\begin{aligned}
I(f)-M_n(f)&=\int_a^b f(x)dx-h\sum_{k=0}^{n-1} f(\bar{x}_k)=\sum_{k=0}^{n-1}\int_{x_k}^{x_{k+1}}\left[f(x)-f(\bar{x}_k)\right]dx\\
&=\sum_{k=0}^{n-1}\int_{x_k}^{x_{k+1}}\left[f'(\bar{x}_k)(x-\bar{x}_k)+\frac{f''(\zeta)}{2}(x-\bar{x}_k)^2\right]dx\\
&=\sum_{k=0}^{n-1}f'(\bar{x}_k)\int_{x_k}^{x_{k+1}}(x-\bar{x}_k)dx+\frac{1}{2}\sum_{k=0}^{n-1}\int_{x_k}^{x_{k+1}}f''(\zeta)(x-\bar{x}_k)^2dx,\\
&=\left[\int_{x_k}^{x_{k+1}}(x-\bar{x}_k)dx=\frac{1}{2}(x-\bar{x}_k)_{x_k}^{x_{k+1}}=0\right]\\
&=\frac{1}{2}\sum_{k=0}^{n-1}\int_{x_k}^{x_{k+1}}f''(\zeta)(x-\bar{x}_k)^2dx.
\end{aligned} \tag{1.11}$$

Since $\left(x-\bar{x}_k\right)^2$ does not change the sign on the interval $\left[x_k, x_{k+1}\right]$ and continuity of the function $f''(\xi)$ on the closed interval $[a,b]$ leads to the mean value theorem for integral.

It state that if $f:[a,b]\to R$ is continuous and $g(x)$ is an integrable function that does not change sign on $[a,b]$, then there exists $c\in(a,b)$ such that

$$\int_a^b f(x)g(x)dx = f(c)\int_a^b g(x)dx\,, \tag{1.12}$$

Using (1.12) we have

$$\int_{x_k}^{x_{k+1}} f''(\zeta)\left(x-\bar{x}_k\right)^2 dx = L\int_{x_k}^{x_{k+1}} \left(x-\bar{x}_k\right)^2 dx = \frac{L}{12}h^3\,,$$

Therefore Eq. (1.11) has the form

$$I(f)-M_n(f)=\frac{L}{24}nh^3=\frac{(b-a)^3}{24}\frac{L}{n^2}$$

For the trapezoidal rule we use linear interpolation formula on the interval $\left[x_k, x_{k+1}\right]$ which is

$$L_1(x)=\frac{x-x_k}{h}f(x_{k+1})+\frac{x_{k+1}-x}{h}f(x_k)\,,$$

then the error term of the function on the interval $\left[x_k, x_{k+1}\right]$ is

$$f(x)-L_1(x)=\frac{1}{2}(x-x_k)(x-x_{k+1})f''(\varsigma)\,. \tag{1.13}$$

Using (1.13) and taking into account the mean value theorem we obtain

$$\begin{aligned} I(f)-T_n(f) &= \int_a^b f(x)dx - h\left[\frac{f(a)}{2}+\sum_{k=1}^{n-1} f(x_k)+\frac{f(b)}{2}\right] \\ &= \sum_{k=0}^{n-1}\int_{x_k}^{x_{k+1}}\left[f(x)-L_1(x)\right]dx = \frac{1}{2}\sum_{k=0}^{n-1}\int_{x_k}^{x_{k+1}}(x-x_k)(x-x_{k+1})f''(\varsigma_x)dx\,, \\ &= \frac{L}{2}\sum_{k=0}^{n-1}\int_{x_k}^{x_{k+1}}(x-x_k)(x-x_{k+1})dx = -\frac{L}{2}n\frac{h^3}{6} = -\frac{(b-a)^3}{12}\frac{L}{n^2}. \end{aligned} \tag{1.14}$$

Example 1: Consider rectangle rules for different type of functions $f(x)$ on the interval $[0,1]$. Let functions f be convex

$$\overline{R_n}=\frac{1}{n}\sum_{k=0}^{n-1} f\left(\frac{k}{n}\right)\approx\int_0^1 f(x)\,dx.\,, \tag{1.15}$$

TABLE 1

	Rectangle Rules			Trapezoidal Rule
n	$f(x)=x$	$f(x)=x^{1/2}$	$f(x)=\sin \pi x$	$f(x)=\sin \pi x$
2	0.25000000	0.35355334	0.50000000	0.50000000
8	0.43750000	0.59563020	0.62841740	0.62841740
32	0.48437500	0.64993387	0.63610828	0.63610828
128	0.49609364	0.66261916	0.63658754	0.63658754
512	0.49902287	0.66567123	0.63661644	0.63661644
2048	0.49975292	0.66641684	0.63661782	0.63661782
4096	0.49987386	0.66653536	0.63661043	0.63661043
Exact value	0.50000000	0.66666667	0.63661782	0.63661977
Error	0.00012614	0.00013134	0.00000195	0.00000934

These examples emphasize the very slow rate of convergence of primitive rules. In the first example $f(x)=x$, the exact error is 1/2*n*. For *n*=4096, the observed error is 0.000126, whereas the exact error should be 0.000122. The difference is due to round-off.

The function $y=x^{1/2}$ does not have a bounded derivative in [0, 1]. For this function, $w(\delta)=\delta^{1/2}$, the error is at worst $n^{-1/2}$. For *n*=4096, this amount to 0.016, but observed error is 0.00013. For the function $f(x)=\sin \pi x,\ f(0)=f(1)$ the results are identical with exact upto 5 digits so that both methods are quite accurate.

In the second place, many of examples have been "over-computed". That is, more work has been done than was necessary to obtain the answer to a given number of significant figures. This reflects not only the conscious desire on our part to exhibit the deleterious effects of round-off but also a frame of mind easily slipped into:

Despite the poor convergence properties of Reimann sums, it must not be ruled out as a practical device. For example, the Reimann sum

$$\sum_{i=1}^{n} f(\zeta_k)(z_i - z_{i-1}) \approx \int_c f(z)dz, \qquad \zeta_k \in (z_{i-1}, z_i),, \tag{1.16}$$

has been used to obtain integrals over arcs in the complex plane. Reimann sums come into their own when the arcs are hard to handle, that is, when they are given by a complicated parametric representation or merely by data points. Primitive situations may require primitive tools.

Reimann sums have also been used in the computation of certain integrals that come up in the solution of partial differential equations over complicated domains by means of the method of least squares. They have been applied to integral equations, too.

1.2 *Higher Rules as Reimann Sums*

It follows from the definition (1.16) that an integration formula

$$\int_a^b f(x)dx \approx \sum_{i=1}^{n} w_i f(\xi_i)$$

will be a Reimann sum if we can find points $a = x_0 < x_1 < ... < x_n = b$ so that

$$x_1 - x_0 = w_1, ..., x_n - x_{n-1} = w_n \text{ and } x_{i-1} \le \xi_i \le x_i .$$

Many higher rules qualify as Reimann sums. Thus the trapezoidal rule, Simpson's rule and the (closed) Newton-Cotes rules of order n=4, 5, 6, 7 are all Reimann sums.

1.3 Simpson's Rule

This rule is very frequently used in obtaining approximate integrals, either in its compound form or as a component in an automatic integration scheme. One of the scholars Milton Abramowitz used to say-somewhat in jest-that 95% of all practical work in numerical analysis boiled down to applications of Simpson's rule and linear interpolation.

Theorem 1.6: Let $f(x) \in C^4[a,b]$ then

$$\begin{aligned} I(f) - S_n(f) &= \int_a^b f(x)dx - \frac{b-a}{6}\left[f(a) + 4f\left(\frac{a+b}{2}\right) + f(b) \right], \\ &= -\frac{(b-a)^5}{2880} f^{(4)}(\xi), \qquad a < \xi < b. \end{aligned} \tag{1.17}$$

for the composite Simpson rule we have

$$\begin{aligned} I(f) - S_n(f) &= \int_a^b f(x)dx - \frac{h}{3}\left[f_0 + 4(f_1 + f_3 + \ldots + f_{2n-1}) + 2(f_2 + f_4 + \ldots + f_{2n-2}) + f_{2n} \right] \\ &= -\frac{(b-a)^5}{2880n^4} f^{(4)}(\xi), \qquad a < \xi < b. \end{aligned} \tag{1.18}$$

The Simpson approximation

$$\int_a^b f(x)dx \approx \frac{b-a}{6}\left[f(a) + 4f\left(\frac{a+b}{2}\right) + f(b) \right], \tag{1.19}$$

is exact for all polynomials of degree three or less.

Let us prove the **exactness of Simpson rule** given in (1.19). In fact, let $P_3(x)$ be polynomial of degree three

$$P_3(x) = a_0 + a_1 x + a_2 x^2 + a^3 x^3 = P_2(x) + a^3 x^3,$$

then integrate it

$$\int_a^b P_3(x)dx = \int_a^b P_2(x)dx + a^3 \int_a^b x^3 dx = \int_a^b P_2(x)dx + \frac{a^3}{4}\left(b^3 - a^3\right), \tag{1.20}$$

It is known that

$$\int_a^b P_2(x)dx = \frac{b-a}{6}\left[P_2(a) + 4P_2\left(\frac{a+b}{2}\right) + P_2(b) \right], \tag{1.21}$$

On the other hand

$$\frac{a_3}{4}\left(b^3 - a^3\right) = \frac{b-a}{6}\left\{ a_3 a^3 + 4a_3\left(\frac{a+b}{2}\right)^2 + a_3 b^3 \right\}, \tag{1.22}$$

From (1.20)-(1.22) it follows that

$$\int_a^b P_3(x)dx \approx \frac{b-a}{6}\left[P_3(a)+4P_3\left(\frac{a+b}{2}\right)+P_3(b)\right]. \tag{1.23}$$

Simpson's rule is most frequently applied in its extended or compound form (Some authors refer to the compound form of a rule as a composite rule.) The interval $[a,b]$ is divided into a number of equal even subintervals and Simpson's rule is applied to each i.e.

Let $a = x_0 < x_1 < \ldots < x_{2n-1} < x_{2n} = b$ be a sequence of equally spaced points in $[a,b]$ and consider subintervals $[x_0,x_2],[x_2,x_4],\ldots,[x_{2n-2},x_{2n}]$ with length $x_{i+1}-x_i = 2h$. Set $f_i = f(x_i)$, then the compound Simpson's rule can be get by applying (1.19) to each interval $[x_{2i-2},x_{2i}]$, $i = 1,\ldots,n$

$$I(f) = \int_a^b f(x)dx = \frac{h}{3}\left[f_0 + 4(f_1+f_3+\ldots+f_{2n-1}) + 2(f_2+f_4+\ldots+f_{2n-2}) + f_{2n}\right] + E_n.\,, \tag{1.24}$$

The remainder E_n is given by

$$E_n = -\frac{nh^5}{90} f^{(4)}(\xi), \quad a < \xi < b.\,, \tag{1.25}$$

If n designates the (even) number of subdivisions of $[a,b]$ then $h = \frac{b-a}{2n}$, so that

$$E_n = -\frac{(b-a)^5}{2880n^4} f^{(4)}(\xi), \quad a < \xi < b.\,, \tag{1.26}$$

For functions that have four continuous derivatives, Simpson's rule (or compound Simpson's rule) converges to the true value of the integral with rapidity N^{-4} at worst. In practise therefore, one might expect the use of 100 subintervals to return a result with four additional correct decimals to that returned using 10 subintervals.

Example 2: Let

$$f_1(x) = \frac{1}{1+x},\ f_2(x) = \frac{x}{e^x-1},\ f_3(x) = x^{3/2},\ f_3(x) = x^{1/2}.$$

n	$f_1(x)$	$f_2(x)$	$f_3(x)$	$f_4(x)$
2	0.69325395	0.77750400	0.40043191	0.65652627
4	0.69315450	0.77750446	0.40007723	0.66307925
8	0.69314759	0.77750450	0.40001368	0.66539813
16	0.69314708	0.77750438	0.40000235	0.66621804
32	0.69314683	0.77750416	0.40000033	0.66650782
64	0.69314670	0.77750411	0.39999984	0.66661024
128	0.69314664	0.77750407	0.39999973	0.66664641
Exact value	0.69314718	0.77750463	0.40000000	0.66666667
Error	0.0000005	0.00000056	0.00000027	0.00002029

The theoretical error bound may be easily computed for the first function. We selected *n=8*. Then from (1.25) we have

$$|E| \le \frac{1}{180\cdot 16^4}\max_{0\le x\le 1}\left|f_1^{(4)}(x)\right|$$

The observed error at n=8 is 0.0000004. Note that after n=16 the accuracy of the answer has deteriorated due to roundoff. Now $f_1^{(4)}(x) = 24(1+x)^{-5}$ so that

$$\max_{0\le x\le 1}\left| f_1^{(4)}(x)\right| = 24 \Rightarrow |E| \le \frac{24}{180\cdot 16^4} \approx 0.000002.$$

The second function is a Debye function. It has only an apparent singularity at *x=0*, and we have $\lim_{x\to 0} f_2(x) = 1$. The value 1 was inserted at *x=0*. Since $f_2(x)$ has derivatives of all orders in [0, 1], we can use error estimate but we have to compute the fourth derivative of $x/(e^x - 1)$ and estimate its maximum value. This is a troublesome computation; we can avoid it by using the series expansion for $x/(e^x - 1)$. We have

$$f_2(x) = \frac{x}{e^x - 1} = \sum_{n=0}^{\infty} \frac{B_n}{n!} x^n,$$

where B_n is the nth Bernoulli number. Hence

$$\left(\frac{x}{e^x - 1}\right)^{(4)} = B_4 + \frac{B_6}{2!}x^2 + \frac{B_8}{4!}x^4 + ...$$

and

$$m = \max_{0\le x\le 1}\left| f_2^{(4)}(x)\right| \le B_4 + \frac{B_6}{2!} + \frac{B_8}{4!} + ... \le 0.05.$$

Selecting *n=2*, we have $|E| \le 0.05/180\cdot 4^4 \approx 10^{-6}$. This high accuracy is borne out by comparison with the exact value.

1.3 Nonequally Spaced Abscissas

The problem here is to obtain the integral $\int_a^b f(x)dx$ when $f(x)$ has been tabulated at nonequally spaced abscissas. This case arises frequently when experimental data are processed but is not likely to arise with functions defined by ”formulas” except when $y = f(x)$ is given in parametric form by the pair of formulas $\{x = x(t), y = y(t) : c \le t \le d\}$ and it is inconvenient to evaluate the inverse function $t = t(x)$.

The following rule, based on overlapping parabolas, is frequently employed. It combines both an integrating and a smoothing feature. Let the abscissas be $a = x_0 < x_1 < ... < x_n = b$ and let

$$P_2(x_{i-1}, x_i, x_{i+1}) = a_i x^2 + b_i x + c_i, \tag{1.27}$$

be the quadratic polynomial that interpolates to $f(x)$ at the three consecutive points $x_{i-1},\ x_i,\ x_{i+1}\ (i = 1, 2, ..., n-1)$ then for $i = 1, 2, ..., n-2$ use the approximation

$$\int_{x_i}^{x_{i+1}} f(x)\,dx \approx \tfrac{1}{2}\int_{x_i}^{x_{i+1}} \left(P_2\left(x_{i-1},x_i,x_{i+1}\right)+P_2\left(x_{i-1},x_i,x_{i+2}\right)\right)dx$$
$$= \frac{a_i+a_{i+1}}{2}\left(\frac{x_{i+1}^3-x_i^3}{3}\right)+\frac{b_i+b_{i+1}}{2}\left(\frac{x_{i+1}^2-x_i^2}{2}\right) \quad , \tag{1.28}$$
$$+\frac{c_i+c_{i+1}}{2}\left(x_{i+1}-x\right).$$

Over the first and last intervals no smoothing is done and the approximations

$$\begin{aligned}\int_{x_0}^{x_1} f(x)\,dx &\approx \int_{x_0}^{x_1} P_2\left(x_0,x_1,x_2\right)dx,\\ \int_{x_{n-1}}^{x_n} f(x)\,dx &\approx \int_{x_{n-1}}^{x_n} P_2\left(x_{n-2},x_{n-1},x_n\right)dx,\end{aligned} \quad , \tag{1.29}$$

are used.

When a or b do not occur at one of the given abscissas but, say, $x_i < a < x_{i+1}$. *then if* $1 \le i \le n-2$, we use the approximation

$$\int_{x_i}^{x_{i+1}} f(x)\,dx \approx \tfrac{1}{2}\int_{x_i}^{x_{i+1}} \left(P_2\left(x_{i-1},x_i,x_{i+1}\right)+P_2\left(x_{i-1},x_i,x_{i+2}\right)\right)dx.\,, \tag{1.30}$$

If $x_0 < a < x_1$ *and also if* $a < x_0$,we use the approximation

$$\int_{a}^{x_1} f(x)\,dx \approx \int_{a}^{x_1} P_2\left(x_0,x_1,x_2\right)\,dx.\,, \tag{1.31}$$

Similar formulas hold for $x_i < b < x_{i+1}$, $1 \le i \le n-2$, $x_{n-1} < b < x_n$, *and* $x_n < b$. Finally, for $x_{n-1} < a$ *or* $b < x_1$, we use formulas similar to (1.30)-(1.31).

1.4 Product Integration Rule and Its Estimation

1.4.1 Product integration rule on J (finite or infinite).

Let us consider product integral of the form

$$I(f)=\int_J k(x)\,f(x)\,dx\,, \tag{1.32}$$

where J is given interval (finite or infinite), $k(x)$ is a weight function and $f(x)$ is given function smooth enough.

Approximating product integral (1.32) by an quadrature formula named interpolatory product integration rule is of the form

$$I_n(k,f)\approx\int_J k(x)\,J_n f(x)\,dx\approx\sum_{j=0}^{n} w_{n,j}(k)\,f(x_{n,j})\,, \tag{1.33}$$

where the points $x_{n,j}$ are fixed and distinct numbers lying in the interval J the function $J_n f$ denotes the Lagrange interpolation polynomial (LIP) of degree at most n which coincides with f at the nodes $\{x_{n,i}\}_{i=0}^{n}$. These nodes are considered the zeros of $p(x)\in W_n$, where W_n is the set of monic (not compulsory) polynomials of degree n, having n distinct zeros on J.

To construct an interpolation quadrature formula for approximate integration (1.32) there are two ways:

(1) Interpolate the function $f(x)$ at the $n+1$ points $\{x_0,x_1,\ldots,x_n\}$ by a polynomial of class $\hat{P}_n$, then integrate the interpolating polynomial and express the result in the form (1.33).

(2) Select the constants $\{w_0,w_1,\ldots,w_n\}$ so that the error

$$E(k,f)=\int_J k(x)\,f(x)\,dx-\sum_{j=0}^{n} w_j(k)\,f(x_j)$$

is zero for $f(x)=\{1,x,x^2,\ldots,x^n\}$.

Theorem 1.7: Two ways of determining $\{w_0,w_1,\ldots,w_n\}$, (1) and (2) yield the same numbers.

Proof: Case (1): Let the polynomial of class $\hat{P}_n$ that interpolates to $f(x)$ at $\{x_0,x_1,\ldots,x_n\}$ is given explicitly by

$$P_n(x)=\sum_{j=0}^{n} f(x_j)\,l_j(x)\,, \tag{1.34}$$

where $l_j(x)=\prod_{\substack{i=0\\ i\neq j}}^{n}\dfrac{x-x_i}{x_j-x_i}=\dfrac{w(x)}{w'(x_j)(x-x_j)}$, $w(x)=(x-x_0)(x-x_1)\ldots(x-x_n)$

Hence the approximate integration formula is

$$\begin{aligned}\int_a^b k(x)\,f(x)\,dx&\approx\int_a^b k(x)\,P_n(x)\,dx\\&=\sum_{j=0}^{n} f(x_j)\int_a^b k(x)\frac{w(x)}{w'(x_j)(x-x_j)}\,dx=\sum_{j=0}^{n} w_j\,f(x_j),\end{aligned} \tag{1.35}$$

where

$$w_j = \int_a^b k(x) \frac{w(x)}{w'(x_j)(x - x_j)} dx, \quad j = 0,1,...,n \,, \tag{1.36}$$

The Eq. (1.35) is the interpolatory quadrature formula and weights is defined by (1.36)

Case (2): Let us construct quadrature formula for the case of (2). Let

$$E(k,1) = 0,\ E(k,x) = 0,\ \ldots,\ E(k,x^n) = 0, \tag{1.37}$$

this leads to

$$\begin{cases} w_0 + w_1 + \ldots + w_n = \int_a^b k(x)dx = m_0 \\ w_0 x_0 + w_1 x_1 + \ldots + w_n x_n = \int_a^b k(x) x dx = m_1 \\ \ldots \\ w_0 x_0^n + w_1 x_1^n + \ldots + w_n x_n^n = \int_a^b k(x) x dx = m_n \end{cases} \tag{1.38}$$

The constants $m_j,\ j = 0,1,...,n$ are the moments of the weighting function $k(x)$. The matrix of this system is the Vandermonde matrix which is nonsingular if the nodes x_i are distinct. Therefore the system (1.38) has one and only one solution for the w_j coefficients.

Now it can be easily shown that if the weights w_j are the solution of (1.38) then it can presented by (1.36) and vice-verse is also true.

1.4.2 **Error Estimation**

It is assumed that $f \in B^n(J)$, the space of functions whose *n*th derivative is continuous and bounded on J , and $k(x) \in X$, where X is the normed space which guarantees existence of the integrals $I(k,f)$ and $I_n(k,f)$.

Bounds for quadrature error (QE) (1.33) for the cases of bounded *J* are obtained in the form

$$|I(k,f) - I_n(k,f)| = \left| \int_J k(x)[f(x) - J_n f(x)] dx \right|$$

$$\leq \|k\|_X \|f - J_n f\|_B \leq M_n(\rho) \|k\|_X \left\| f^{(n)} \right\|_\infty, \tag{1.39}$$

where $M_n(\rho)$ is independent of k and f and is smallest constant such that (1.39) is valid for all $k \in X$ and $f \in B^n(J)$.

From the derived expression (1.39) for $M_n(\rho)$ it is then possible to calculate the optimal nodes as the zeros of p^* where

$$M_n(p) = \inf_{\rho \in W_n} M_n(p).$$

In some cases, we are only able to obtain asymptotic expressions for $M_n(p)$ but from such expressions useful statements can be made on the relative preferability of different sets of nodes.

Several authors, e.g., [5], [7], [9], [14], [16]-[19] have considered convergence of product integration rules over finite and infinite intervals, but generally their results have been given in terms of orders of convergence without particular attention to obtaining sharp error bounds.

Product integration over $J = [-1,1]$.

In this section, we establish the basic procedure by considering the finite interval $J = [-1,1]$ and the function space $X = L_p[-1,1]$, and derive useful results when the nodes are the zeros of a Jacobi polynomials particularly Chebyshev polynomials. Throughout the paper we assume that p and q as a pair of positive numbers satisfying

$$\frac{1}{p} + \frac{1}{q} = 1 \tag{1.40}$$

It is known that for the finite interval $J = [-1,1]$ case and when $k \in L_p[-1,1]$, $\rho \geq 1$ and $f \in C[-1,1]$ then $I(k,f)$ exists and quadrature error (QE) is defined by

$$E_n(k,f) = I(k,f) - I_n(k,f) = \int_{-1}^{1} k(x)[f(x) - J_n f(x)]\, dx. \tag{1.41}$$

Hence for any $k \in L_p, p \geq 1$ according to error term of (1.39) we have

$$|E_n(k,f)| \leq \|k\|_p \|f - J_n f\|_q, \tag{1.42}$$

and this estimation is called best bound in the sense that for any $f \in C$, the quantity $\|f - J_n f\|_q$ can't be replaced by anything smaller than (1.42) for all $k \in L_p$.

Now suppose $f \in C^r[-1,1]$, $r \leq n$, then $\|f - J_n f\|_q$ is bounded by a best estimate of the form

$$\|f - J_n f\|_q \leq M_n(\rho,p) \left\|f^{(r)}\right\|_\infty, \tag{1.43}$$

where the zeros of $\rho \in W_n$ are the nodes of interpolation. Thus for $f \in C^r[-1,1]$ we may write

$$|E_n(k,f)| \leq M_n(\rho,p) \|k\|_p \left\|f^{(r)}\right\|_\infty, \tag{1.44}$$

which is valid for all $k \in L_p$.

From (1.41) it follows that

$$|E_n(k,f)| = \frac{|I(k,f) - I_n(k,f)|}{\|k\|_p \left\|f^{(r)}\right\|_\infty} \|k\|_p \left\|f^{(r)}\right\|_\infty, \tag{1.45}$$

and comparing with (1.44), we obtain

$$M_n(\rho,p) = \sup_{f \in C^r} \sup_{k \in L_p} \frac{|I(k,f) - I_n(k,f)|}{\|k\|_p \left\|f^{(r)}\right\|_\infty} = \sup_{f \in C^r} \frac{\|f - J_n f\|_q}{\left\|f^{(r)}\right\|_\infty} \tag{1.46}$$

The next step is to choose an optimal set of nodes or equivalently a monic polynomial ρ^*, so that $M_n(\rho,p)$ is minimized with $\rho = \rho^*$, i.e.

$$M_n^*(p^*,p) = \inf_{\rho\in W_n} \sup_{f\in C^r} \frac{\|f-J_nf\|_q}{\|f^{(r)}\|_\infty}. \tag{1.47}$$

We observe that for $r = 0$ and $p = 1$, the optimal choice of nodes is precisely that for best uniform interpolatory approximation, that is the classical problem of interpolation.

Primary concern is to minimize $M_n^*(p^*,p)$ defined by (2.8) in differentiable class of functions $f(x)$.

Theorem 1: For $k \in L_p$ and $f \in C^n$ the error for interpolatory product integration based on the zeros of $\rho(x) \in W_n$ has a best bound given by

$$|E_n(k,f)| \le M_n(\rho,p)\|k\|_p\|f^{(r)}\|_\infty = \frac{1}{n!}\|\rho\|_q\|k\|_p\|f^{(n)}\|_\infty$$

Furthermore, for the quadrature rule based on the optimal nodes which minimize $M_n(\rho,p)$ is

$$|E_n^*(k,f)| \le M_n^*(\rho,p)\|k\|_p\|f^{(r)}\|_\infty = \frac{1}{n!}\|\rho *\|_q\|k\|_p\|f^{(n)}\|_\infty$$

$$= \frac{1}{n!}\inf_{\rho\in W_n}\|\rho\|_q\,\|k\|_p\|f^{(n)}\|_\infty.$$

Proof: For $f \in C^n$ we have

$$f(x) - J_nf(x) = \frac{1}{n!}\rho(x)f^{(n)}(\theta_x),\ \ \theta_x \in (-1,1). \tag{1.48}$$

To prove (1.48) let $-1 \le x_1 < x_2 < \cdots < x_n \le 1$ and $J_nf(x)$ be Lagrange polynomials of degree at most $n-1$ in the form

$$J_nf(x) = f(x_1)L_{n,1}(x) + f(x_2)L_{n,2}(x) + \cdots + f(x_n)L_{n,n}(x)$$

$$= \sum_{j=1}^n f(x_j)\,L_{n,j}(x), L_{n,j}(x) = \prod_{\substack{i=1\\ i\ne j}}^n \frac{x-x_i}{x_j-x_i}.$$

Note that if $x = x_k, k = 1,2,\dots,n$ then

$$f(x_k) = P(x_k),\ \ k = 1,2\dots,n \tag{1.49}$$

And choosing $\theta_x \in (-1,1)$ yields the result

$$f(x) = P(x) + \frac{f^{(n)}(\theta_x)}{n!}(x-x_1)(x-x_2)\dots(x-x_n). \tag{1.50}$$

If $x \ne x_k, k = 1,2,\dots,n$ define the function $g(t)$ for $t \in [-1,1]$ by

$$g(t) = f(t) - P(t) - [f(x) - P(x)]\frac{(t-x_1)(t-x_2)\dots(t-x_n)}{(x-x_1)(x-x_2)\dots(x-x_n)}$$

$$= f(t) - P(t) - [f(x) - P(x)]\prod_{i=1}^n \frac{(t-x_i)}{(x-x_i)} \tag{1.51}$$

Since $f \in C^{(n)}[-1,1]$ and $P \in C^\infty[-1,1]$ it follows that $g \in C^{(n)}[-1,1]$. For $t = x_k$ we have

$$g(x_k) = f(x_k) - P(x_k) - [f(x) - P(x)]\prod_{i=1}^n \frac{(x_k-x_i)}{(x-x_i)}$$

$$= 0 - [f(x) - P(x)]0 = 0, k = 1,2,\dots,n \tag{1.52}$$

And

$$g(x) = f(x) - P(x) - [f(x) - P(x)] \cdot 1 = 0 \tag{1.53}$$

Since $g \in C^{(n)}[-1,1]$ and g is zero at n distinct numbers $\{x, x_2, \dots, x_n\}$ by generalized Roll's Theorem there exists a number $\theta_x \in [-1,1]$ for which

$$g^{(n)}(\theta_x) = 0, \tag{1.54}$$

On the other hand

$$0 = g^{(n)}(\theta_x) = f^{(n)}(\theta_x) - P^{(n)}(\theta_x) - [f(x) - P(x)] \frac{d^n}{dt^n}\left[\prod_{i=1}^{n} \frac{(t-x_i)}{(x-x_i)}\right]_{t=\theta} \tag{1.55}$$

Since $P(x)$ is a polynomial of degree at most $n-1$ so that $P^{(n)}(x) = 0$. Also

$$\frac{d^n}{dt^n}\prod_{i=1}^{n} \frac{(t-x_i)}{(x-x_i)} = \frac{d^n}{dt^n}\left\{\frac{1}{\prod_{i=1}^{n}(x-x_i)} t^n + lower\ degree\ terms\ of\ t\right\} = \frac{n!}{\prod_{i=1}^{n}(x-x_i)} \tag{1.56}$$

Thus, from (1.54)-(1.56) it follows that

$$0 = f^{(n)}(\theta_x) - 0 - [f(x) - P(x)] \frac{n!}{\prod_{i=1}^{n}(x - x_i)}$$

which implies

$$f(x) = P(x) + \frac{f^{(n)}(\theta_x)}{n!}\prod_{i=1}^{n}(x - x_i) = P(x) + \frac{f^{(n)}(\theta_x)}{n!}\rho(x). \tag{1.57}$$

It completes proof of Eq. (1.47). From (1.45) it follows that

$$M_n(\rho, p) = \sup_{f \in C^r} \frac{\|f - J_n f\|_q}{\|f^{(n)}\|_\infty} = \frac{1}{n!}\sup_{f \in C^r} \frac{\|\rho f^{(n)}(\theta_x)\|_q}{\|f^{(r)}\|_\infty} = \frac{1}{n!}\|\rho\|_q \tag{1.58}$$

and the supremum is attained, for example, when $f^{(n)}$ is a constant function. The results for optimal nodes follow immediately from the error term $|E_n(k, f)|$ of the first term. □

In Table 1 we list expressions for the norms $\|\rho\|_1$, $\|\rho\|_2$, and $\|\rho\|_\infty$ for several sets of commonly used nodes $\{x_{n,i}\}_1^n$ (set requires $n \geq 2$), a few of the results are given in [10], [15] and [20], but all can be obtained by a variety of combinations of standard methods.

TABLE 1

Note. $n \geqq 1$ for first three cases, $n \geqq 2$ for last two cases.

nodes $\{x_{n,i}\}_i^n$	$\|\rho\|_1$ $(p=\infty)$	$\|\rho\|_2$ $(p=2)$	$\|\rho\|_\infty$ $(p=1)$
zeros of $T_n(x)$	$\frac{n \operatorname{cosec}(\pi/2n) - 1}{(n^2-1)2^{n-2}}$	$\left(\frac{4n^2-2}{4n^2-1}\right)^{1/2} 2^{1-n} \sim 2^{1-n}$	2^{1-n} OPT
(Chebyshev pts.)	$\sim \pi^{-1} 2^{3-n}$		
zeros of $P_n(x)$	$\frac{2^{n+1}(n!)^2}{(n+1)(2n)!} \sum_{i=1}^{n} \lvert P_{n-1}(x_{n,i})\rvert$	$\frac{2^{n+1/2}(n!)^2}{(2n+1)^{1/2}(2n)!}$ OPT	$\frac{2^n(n!)^2}{(2n)!} \sim (\pi n)^{1/2} 2^{-n}$
(Legendre pts.)	$\sim \frac{(\Gamma(3/4))^2 2^{3-n}}{\pi^{3/2}}$	$\sim \pi^{1/2} 2^{-n}$	

zeros of $U_n(x)$	2^{1-n} OPT	$2^{-n}(\sum_{k=0}^{n} 1/(k+(1/2)))^{1/2}$ $\sim \frac{(\gamma+\ln(4n+6))^{1/2}}{2^n}$	$(n+1)2^{-n}$
zeros of $T_n(x)-T_{n-2}(x)$	2^{2-n}	$2^{2-n}\left(\frac{8}{3}-\frac{3\Gamma(n-\frac{5}{2})}{2\Gamma(n+\frac{3}{2})}\right)^{1/2}$	2^{2-n} n even
or $(1-x^2)U_{n-2}(x)$		$\sim 3^{-1/3}2^{7/2-n}$	$<2^{2-n}$, n odd
(Clenshaw-Curtis pts.)			$\sim 2^{2-n}$
zeros of $T_n(x\cos\pi/2n)$	$\frac{n\sec^n(\pi/2n)\cot(\pi/2n)}{(n^2-1)2^{n-2}}$	$\frac{n\sec^n(\pi/2n)}{(4n^2-1)^{1/2}2^{n-2}}\sim 2^{1-n}$	$2^{1-n}\sec^n(\pi/2n)$
(extended Chebyshev pts.)	$\sim \pi^{-1}2^{3-n}$		$\sim 2^{1-n}$