Chapter 3
Orthogonal Polynomials and Related Approximation Results

The Fourier spectral method is only appropriate for problems with periodic boundary conditions. If a Fourier method is applied to a non-periodic problem, it inevitably induces the so-called Gibbs phenomenon, and reduces the global convergence rate to $O(N^{-1})$ (cf. Gottlieb and Orszag (1977)). Consequently, one should not apply a Fourier method to problems with non-periodic boundary conditions. Instead, one should use orthogonal polynomials which are eigenfunctions of some singular Sturm-Liouville problems. The commonly used orthogonal polynomials include the Legendre, Chebyshev, Hermite and Laguerre polynomials.

The aim of this chapter is to present essential properties and fundamental approximation results related to orthogonal polynomials. These results serve as preparations for polynomial-based spectral methods in the forthcoming chapters. This chapter is organized as follows. In the first section, we present relevant properties of general orthogonal polynomials, and set up a general framework for the study of orthogonal polynomials. We then study the Jacobi polynomials in Sect. 3.2, Legendre polynomials in Sect. 3.3 and Chebyshev polynomials in Sect. 3.4. In Sect. 3.5, we present some general approximation results related to these families of orthogonal polynomials. We refer to Szegö (1975), Davis and Rabinowitz (1984) and Gautschi (2004) for other aspects of orthogonal polynomials.

3.1 Orthogonal Polynomials

Orthogonal polynomials play the most important role in spectral methods, so it is necessary to have a thorough study of their relevant properties. Our starting point is the generation of orthogonal polynomials by a three-term recurrence relation, which leads to some very useful formulas such as the Christoffel-Darboux formula. We then review some results on zeros of orthogonal polynomials, and present efficient algorithms for their computations. We also devote several sections to discussing some important topics such as Gauss-type quadrature formulas, polynomial interpolations, discrete transforms, and spectral differentiation techniques.

J. Shen et al., *Spectral Methods: Algorithms, Analysis and Applications*,
Springer Series in Computational Mathematics 41, DOI 10.1007/978-3-540-71041-7_3,

3.1.1 Existence and Uniqueness

Given an open interval $I := (a,b)$ $(-\infty \le a < b \le +\infty)$, and a generic weight function ω such that

$$\omega(x) > 0,\ \forall x \in I \ \text{ and } \ \omega \in L^1(I), \tag{3.1}$$

two functions f and g are said to be *orthogonal* to each other in $L^2_\omega(a,b)$ or orthogonal with respect to ω if

$$(f,g)_\omega := \int_a^b f(x)g(x)\omega(x)dx = 0.$$

An algebraic polynomial of degree n is denoted by

$$p_n(x) = k_n x^n + k_{n-1}x^{n-1} + \ldots + k_1 x + k_0, \quad k_n \neq 0, \tag{3.2}$$

where $\{k_i\}$ are real constants, and k_n is the leading coefficient of p_n.

A sequence of polynomials $\{p_n\}_{n=0}^\infty$ with $\deg(p_n) = n$ is said to be *orthogonal* in $L^2_\omega(a,b)$ if

$$(p_n, p_m)_\omega = \int_a^b p_n(x)p_m(x)\omega(x)dx = \gamma_n \delta_{mn}, \tag{3.3}$$

where the constant $\gamma_n = \|p_n\|^2_\omega$ is nonzero, and δ_{mn} is the Kronecker delta.

Throughout this section, $\{p_n\}$ denotes a sequence of polynomials orthogonal with respect to the weight function ω, and p_n is of degree n.

Denote by P_n the set of all algebraic polynomials of degree $\le n$, namely,

$$P_n := \text{span}\left\{1, x, x^2, \ldots, x^n\right\}. \tag{3.4}$$

By a dimension argument,

$$P_n = \text{span}\{p_0, p_1, \ldots, p_n\}. \tag{3.5}$$

A direct consequence is the following.

Lemma 3.1. p_{n+1} *is orthogonal to any polynomial* $q \in P_n$.

Proof. Thanks to (3.5), for any $q \in P_n$, we can write

$$q = b_n p_n + b_{n-1} p_{n-1} + \ldots + b_0 p_0.$$

Hence,

$$(p_{n+1}, q)_\omega = b_n(p_{n+1}, p_n)_\omega + b_{n-1}(p_{n+1}, p_{n-1})_\omega + \ldots + b_0(p_{n+1}, p_0)_\omega = 0,$$

where we have used the orthogonality (3.3). □

Hereafter, we denote the *monic* polynomial corresponding to p_n by

$$\bar{p}_n(x) := p_n(x)/k_n = x^n + a^{(n)}_{n-1}x^{n-1} + \ldots + a^{(n)}_0. \tag{3.6}$$

We show below that for any given weight function $\omega(x)$ defined in (a,b), there exists a unique family of monic orthogonal polynomials generated by a three-term recurrence formula.

Theorem 3.1. *For any given positive weight function $\omega \in L^1(I)$, there exists a unique sequence of monic orthogonal polynomials $\{\bar{p}_n\}$ with $\deg(\bar{p}_n) = n$, which can be constructed as follows*

$$\begin{aligned} &\bar{p}_0 = 1, \;\; \bar{p}_1 = x - \alpha_0, \\ &\bar{p}_{n+1} = (x-\alpha_n)\bar{p}_n - \beta_n \bar{p}_{n-1}, \quad n \ge 1, \end{aligned} \tag{3.7}$$

where

$$\alpha_n = \frac{(x\bar{p}_n, \bar{p}_n)_\omega}{\|\bar{p}_n\|_\omega^2}, \quad n \ge 0, \tag{3.8a}$$

$$\beta_n = \frac{\|\bar{p}_n\|_\omega^2}{\|\bar{p}_{n-1}\|_\omega^2}, \quad n \ge 1. \tag{3.8b}$$

Proof. It is clear that the first two polynomials are

$$\bar{p}_0(x) \equiv 1, \quad \bar{p}_1(x) = x - \alpha_0.$$

To determine α_0, we see that $(\bar{p}_0, \bar{p}_1)_\omega = 0$ if and only if

$$\alpha_0 = \int_a^b \omega(x) x dx \Big/ \int_a^b \omega(x) dx = \frac{(x\bar{p}_0, \bar{p}_0)_\omega}{\|\bar{p}_0\|_\omega^2},$$

where the denominator is positive due to (3.1).

We proceed with the proof by using an induction argument. Assuming that by a similar construction, we have derived a sequence of monic orthogonal polynomials $\{\bar{p}_k\}_{k=0}^n$. Next, we seek $\bar{p}_{n+1}$ of the form

$$\bar{p}_{n+1} = x\bar{p}_n - \alpha_n \bar{p}_n - \beta_n \bar{p}_{n-1} - \sum_{k=0}^{n-2} \gamma_k^{(n)} \bar{p}_k, \tag{3.9}$$

with α_n and β_n given by (3.8), and we require

$$\left(\bar{p}_{n+1}, \bar{p}_k\right)_\omega = 0, \quad 0 \le k \le n. \tag{3.10}$$

Taking the inner product with $\bar{p}_k$ on both sides of (3.9), and using the orthogonality of $\{\bar{p}_k\}_{k=0}^n$, we find that (3.10) is fulfilled if and only if

$$\begin{aligned} &(\bar{p}_{n+1}, \bar{p}_n)_\omega = (x\bar{p}_n, \bar{p}_n)_\omega - \alpha_n (\bar{p}_n, \bar{p}_n)_\omega = 0, \\ &(\bar{p}_{n+1}, \bar{p}_{n-1})_\omega = (x\bar{p}_n, \bar{p}_{n-1})_\omega - \beta_n (\bar{p}_{n-1}, \bar{p}_{n-1})_\omega = 0, \\ &(\bar{p}_{n+1}, \bar{p}_j)_\omega = (x\bar{p}_n, \bar{p}_j)_\omega - \sum_{k=0}^{n-2} \gamma_k^{(n)} (\bar{p}_k, \bar{p}_j)_\omega \\ &\qquad\qquad \overset{(3.3)}{=} (x\bar{p}_n, \bar{p}_j)_\omega - \gamma_j^{(n)} \|\bar{p}_j\|_\omega^2 = 0, \quad 0 \le j \le n-2. \end{aligned} \tag{3.11}$$

Hence, the first equality implies (3.8a), and by the second one,

$$\beta_n = \frac{(x\bar{p}_n, \bar{p}_{n-1})_\omega}{\|\bar{p}_{n-1}\|_\omega^2} = \frac{(\bar{p}_n, x\bar{p}_{n-1})_\omega}{\|\bar{p}_{n-1}\|_\omega^2} = \frac{\|\bar{p}_n\|_\omega^2}{\|\bar{p}_{n-1}\|_\omega^2},$$

where we have used the fact

$$x\bar{p}_{n-1} = \bar{p}_n + \sum_{k=0}^{n-1} \delta_k^{(n)} \bar{p}_k,$$

and the orthogonality of $\{\bar{p}_k\}_{k=0}^n$ to deduce the last identity. It remains to show that the coefficients $\{\gamma_k^{(n)}\}_{k=0}^{n-2}$ in (3.9) are all zero. Indeed, we derive from Lemma 3.1 that

$$\big(x\bar{p}_n, \bar{p}_j\big)_\omega = (\bar{p}_n, x\bar{p}_j)_\omega = 0, \quad 0 \le j \le n-2,$$

which, together with the last equation of (3.11), implies $\gamma_k^{(n)} \equiv 0$ for $0 \le k \le n-2$, in (3.9). This completes the induction.

Next, we show that the polynomial sequence generated by (3.7)–(3.8) is unique. For this purpose, we assume that $\{\bar{q}_n\}_{n=0}^\infty$ is another sequence of monic orthogonal polynomials. Since $\bar{p}_n$, given by (3.7), is also monic, we have $\deg\big(\bar{p}_{n+1} - \bar{q}_{n+1}\big) \le n$. By Lemma 3.1,

$$(\bar{p}_{n+1}, \bar{p}_{n+1} - \bar{q}_{n+1})_\omega = 0, \quad (\bar{q}_{n+1}, \bar{p}_{n+1} - \bar{q}_{n+1})_\omega = 0,$$

which implies

$$\big(\bar{p}_{n+1} - \bar{q}_{n+1}, \bar{p}_{n+1} - \bar{q}_{n+1}\big)_\omega = 0 \quad \Rightarrow \quad \bar{p}_{n+1}(x) - \bar{q}_{n+1}(x) \equiv 0.$$

This proves the uniqueness. □

The above theorem provides a general three-term recurrence relation to construct orthogonal polynomials, and the constants α_n and β_n can be evaluated explicitly for the commonly used families. The three-term recurrence relation (3.7) is essential for deriving other properties of orthogonal polynomials. For convenience, we first extend it to the orthogonal polynomials $\{p_n\}$, which are not necessarily monic.

Corollary 3.1. *Let $\{p_n\}$ be a sequence of orthogonal polynomials with the leading coefficient $k_n \neq 0$. Then we have*

$$p_{n+1} = (a_n x - b_n)p_n - c_n p_{n-1}, \qquad n \ge 0, \tag{3.12}$$

with $p_{-1} := 0$, $p_0 = k_0$ and

$$a_n = \frac{k_{n+1}}{k_n}, \tag{3.13a}$$

$$b_n = \frac{k_{n+1}}{k_n} \frac{(xp_n, p_n)_\omega}{\|p_n\|_\omega^2}, \tag{3.13b}$$

$$c_n = \frac{k_{n-1}k_{n+1}}{k_n^2} \frac{\|p_n\|_\omega^2}{\|p_{n-1}\|_\omega^2}. \tag{3.13c}$$

Proof. This result follows directly from Theorem 3.1 by inserting $\bar{p}_l = p_l/k_l$ with $l = n-1, n, n+1$ into (3.7) and (3.8). □

The orthogonal polynomials $\{p_n\}$ with leading coefficients $\{k_n\}$ are uniquely determined by (3.12)–(3.13). Interestingly, the following result, which can be viewed as the converse of Corollary 3.1, also holds. We leave its proof as an exercise (see Problem 3.1).

Corollary 3.2. *Let $\{k_n \neq 0\}$ be a sequence of real numbers. The three-term recurrence relation (3.12)–(3.13) generates a sequence of polynomials satisfying the properties:*

- *the leading coefficient of p_n is k_n and $deg(p_n) = n$;*
- *$\{p_n\}$ are orthogonal with respect to the weight function $\omega(x)$;*
- *the L^2_ω-norm of p_n is given by*

$$\gamma_n = \|p_n\|^2_\omega = (a_0/a_n)c_1 c_2 \ldots c_n \gamma_0, \quad n \geq 0, \tag{3.14}$$

where $\gamma_0 = k_0^2 \int_a^b \omega(x)dx$.

An important consequence of the three-term recurrence formula (3.12)–(3.13) is the well-known *Christoff-Darboux formula*.

Corollary 3.3. *Let $\{p_n\}$ be a sequence of orthogonal polynomials with $deg(p_n) = n$. Then,*

$$\frac{p_{n+1}(x)p_n(y) - p_n(x)p_{n+1}(y)}{x-y} = \frac{k_{n+1}}{k_n} \sum_{j=0}^{n} \frac{\|p_n\|^2_\omega}{\|p_j\|^2_\omega} p_j(x)p_j(y), \tag{3.15}$$

and

$$p'_{n+1}(x)p_n(x) - p'_n(x)p_{n+1}(x) = \frac{k_{n+1}}{k_n} \sum_{j=0}^{n} \frac{\|p_n\|^2_\omega}{\|p_j\|^2_\omega} p_j^2(x). \tag{3.16}$$

Proof. We first prove (3.15). By Corollary 3.1,

$$\begin{aligned}
&p_{j+1}(x)p_j(y) - p_j(x)p_{j+1}(y) \\
&\quad = [(a_j x - b_j)p_j(x) - c_j p_{j-1}(x)]p_j(y) \\
&\qquad - p_j(x)[(a_j y - b_j)p_j(y) - c_j p_{j-1}(y)] \\
&\quad = a_j(x-y)p_j(x)p_j(y) + c_j[p_j(x)p_{j-1}(y) - p_{j-1}(x)p_j(y)].
\end{aligned}$$

Thus, by (3.13),

$$\begin{aligned}
&\frac{k_j}{k_{j+1}\|p_j\|^2_\omega} \frac{p_{j+1}(x)p_j(y) - p_j(x)p_{j+1}(y)}{x-y} \\
&\quad - \frac{k_{j-1}}{k_j\|p_{j-1}\|^2_\omega} \frac{p_j(x)p_{j-1}(y) - p_{j-1}(x)p_j(y)}{x-y} = \frac{1}{\|p_j\|^2_\omega} p_j(x)p_j(y).
\end{aligned}$$

This relation also holds for $j = 0$ by defining $p_{-1} := 0$. Summing the above identities for $0 \leq j \leq n$ leads to (3.15).

To prove (3.16), we observe that

$$\begin{aligned}&\frac{p_{n+1}(x)p_n(y)-p_n(x)p_{n+1}(y)}{x-y}\\&\qquad=\frac{p_{n+1}(x)-p_{n+1}(y)}{x-y}p_n(y)-\frac{p_n(x)-p_n(y)}{x-y}p_{n+1}(y).\end{aligned}$$

Consequently, letting $y\to x$, we obtain (3.16) from (3.15) and the definition of the derivative. □

Define the kernel

$$K_n(x,y)=\sum_{j=0}^{n}\frac{p_j(x)p_j(y)}{\|p_j\|_\omega^2}. \tag{3.17}$$

The Christoff-Darboux formula (3.15) can be rewritten as

$$K_n(x,y)=\frac{k_n}{k_{n+1}\|p_n\|_\omega^2}\,\frac{p_{n+1}(x)p_n(y)-p_n(x)p_{n+1}(y)}{x-y}. \tag{3.18}$$

A remarkable property of $\{K_n\}$ is stated in the following lemma.

Lemma 3.2. *There holds the integral equation*:

$$q(x)=\int_a^b q(t)K_n(x,t)\omega(t)dt,\quad \forall q\in P_n. \tag{3.19}$$

Moreover, the polynomial sequence $\{K_n(x,a)\}$ (*resp.* $\{K_n(x,b)\}$) *is orthogonal with respect to the weight function* $(x-a)\omega$ (*resp.* $(b-x)\omega$).

Proof. Thanks to (3.5), for any $q\in P_n$, we can write

$$q(x)=\sum_{j=0}^{n}\hat{q}_j p_j(x)\quad\text{with}\quad \hat{q}_j=\frac{1}{\|p_j\|_\omega^2}\int_a^b q(t)p_j(t)\omega(t)dt.$$

Thus, by definition (3.17),

$$q(x)=\sum_{j=0}^{n}\frac{1}{\|p_j\|_\omega^2}\int_a^b q(t)p_j(x)p_j(t)\omega(t)dt=\int_a^b q(t)K_n(x,t)\omega(t)dt,$$

which gives (3.19).

Next, taking $x=a$ and $q(t)=(t-a)r(t)$ for any $r\in P_{n-1}$ in (3.19) yields

$$0=q(a)=\int_a^b K_n(t,a)r(t)(t-a)\omega(t)dt,\quad \forall r\in P_{n-1},$$

which implies $\{K_n(x,a)\}$ is orthogonal with respect to $(x-a)\omega$.

Similarly, taking $x=b$ and $q(t)=(b-t)r(t)$ for any $r\in P_{n-1}$ in (3.19), we can show that $\{K_n(x,b)\}$ is orthogonal with respect to $(b-x)\omega$. □

3.1.2 Zeros of Orthogonal Polynomials

Zeros of orthogonal polynomials play a fundamental role in spectral methods. For example, they are chosen as the nodes of Gauss-type quadratures, and used to generate computational grids for spectral methods. Therefore, it is useful to derive their essential properties.

Again, let $\{p_n\}$ (with $\deg(p_n) = n$) be a sequence of polynomials orthogonal with respect to the weight function $\omega(x)$ in (a,b). The first important result about the zeros of orthogonal polynomials is the following:

Theorem 3.2. *The zeros of p_{n+1} are all real, simple, and lie in the interval (a,b).*

Proof. We first show that the zeros of p_{n+1} are all real. Assuming $\alpha \pm \mathrm{i}\beta$ are a pair of complex roots of p_{n+1}. Then $p_{n+1}/((x-\alpha)^2+\beta^2) \in P_{n-1}$, and by Lemma 3.1,

$$0 = \int_a^b p_{n+1} \frac{p_{n+1}}{(x-\alpha)^2+\beta^2}\omega dx = \int_a^b \big((x-\alpha)^2+\beta^2\big)\left|\frac{p_{n+1}}{(x-\alpha)^2+\beta^2}\right|^2 \omega dx,$$

which implies that $\beta = 0$. Hence, all zeros of p_{n+1} must be real.

Next, we prove that the $n+1$ zeros of p_{n+1} are simple, and lie in the interval (a,b). By the orthogonality,

$$\int_a^b p_{n+1}(x)\omega(x)dx = 0, \quad \forall n \geq 0,$$

there exists at least one zero of p_{n+1} in (a,b). In other words, $p_{n+1}(x)$ must change sign in (a,b). Let $x_0, x_1, \ldots, x_k$ be all such points in (a,b) at which $p_{n+1}(x)$ changes sign. If $k = n$, we are done, since $\{x_i\}_{i=0}^n$ are the $n+1$ simple real zeros of p_{n+1}. If $k < n$, we consider the polynomial

$$q(x) = (x-x_0)(x-x_1)\ldots(x-x_k).$$

Since $\deg(q) = k+1 < n+1$, by orthogonality, we derive

$$(p_{n+1}, q)_\omega = 0.$$

However, $p_{n+1}(x)q(x)$ cannot change sign on (a,b), since each sign change in $p_{n+1}(x)$ is canceled by a corresponding sign change in $q(x)$. It follows that

$$(p_{n+1}, q)_\omega \neq 0,$$

which leads to a contradiction. □

Another important property is known as the separation theorem.

Theorem 3.3. *Let $x_0 = a$, $x_{n+1} = b$ and $x_1 < x_2 < \ldots < x_n$ be the zeros of p_n. Then there exists exactly one zero of p_{n+1} in each subinterval (x_j, x_{j+1}), $j = 0, 1, \ldots, n$.*

Proof. It is obvious that the location of zeros is invariant with any nonzero constant multiple of p_n and p_{n+1}, so we assume that the leading coefficients $k_n, k_{n+1} > 0$.

We first show that each of the interior subintervals (x_j, x_{j+1}), $j = 1, 2, \ldots, n-1$, contains at least one zero of p_{n+1}, which is equivalent to proving

$$p_{n+1}(x_j)p_{n+1}(x_{j+1}) < 0, \quad 1 \le j \le n-1. \tag{3.20}$$

Since p_n can be written as

$$p_n(x) = k_n(x-x_1)(x-x_2)\ldots(x-x_n),$$

a direct calculation leads to

$$p_n'(x_j) = k_n \prod_{l=1}^{j-1}(x_j - x_l) \cdot \prod_{l=j+1}^{n}(x_j - x_l). \tag{3.21}$$

This implies

$$p_n'(x_j)p_n'(x_{j+1}) = D_{n,j} \times (-1)^{n-j} \times (-1)^{n-j-1} < 0, \tag{3.22}$$

where $D_{n,j}$ is a positive constant. On the other hand, using the facts that $p_n(x_j) = p_n(x_{j+1}) = 0$ and $k_n, k_{n+1} > 0$, we find from (3.16) that

$$-p_n'(x_j)p_{n+1}(x_j) > 0, \quad -p_n'(x_{j+1})p_{n+1}(x_{j+1}) > 0. \tag{3.23}$$

Consequently,

$$\left[p_n'(x_j)p_n'(x_{j+1})\right]\left[p_{n+1}(x_j)p_{n+1}(x_{j+1})\right] > 0 \overset{(3.22)}{\Longrightarrow} p_{n+1}(x_j)p_{n+1}(x_{j+1}) < 0.$$

Next, we prove that there exists at least one zero of p_{n+1} in each of the boundary subintervals (x_n, b) and (a, x_1). Since $p_n(x_n) = 0$ and $p_n'(x_n) > 0$ (cf. (3.21)), the use of (3.16) again gives $p_{n+1}(x_n) < 0$. On the other hand, due to $k_{n+1} > 0$, we have $p_{n+1}(b) > 0$. Therefore, $p_{n+1}(x_n)p_{n+1}(b) < 0$, which implies (x_n, b) contains at least one zero of p_{n+1}. The existence of at least one zero of p_{n+1} in (a, x_1) can be justified in a similar fashion.

In summary, we have shown that each of the $n+1$ subintervals (x_j, x_{j+1}), $0 \le j \le n$, contains at least one zero of p_{n+1}. By Theorem 3.2, p_{n+1} has exactly $n+1$ real zeros, so each subinterval contains exactly one zero of p_{n+1}. □

A direct consequence of (3.22) is the following.

Corollary 3.4. *Let* $\{p_n\}$ *be a sequence of orthogonal polynomials with* $\deg(p_n) = n$. *Then the zeros of* p_n' *are real and simple, and there exists exactly one zero of* p_n' *between two consecutive zeros of* p_n.

3.1.3 Computation of Zeros of Orthogonal Polynomials

We present below two efficient algorithms for computing zeros of orthogonal polynomials.

The first approach is the so-called *Eigenvalue Method.*

Theorem 3.4. *The zeros $\{x_j\}_{j=0}^n$ of the orthogonal polynomial $p_{n+1}(x)$ are eigenvalues of the following symmetric tridiagonal matrix:*

$$A_{n+1} = \begin{bmatrix} \alpha_0 & \beta_1 & & & \\ \beta_1 & \alpha_1 & \beta_2 & & \\ & \ddots & \ddots & \ddots & \\ & & \beta_{n-1} & \alpha_{n-1} & \beta_n \\ & & & \beta_n & \alpha_n \end{bmatrix}, \tag{3.24}$$

where

$$\alpha_j = \frac{b_j}{a_j}, \quad j \geq 0; \quad \beta_j = \frac{1}{a_{j-1}}\sqrt{\frac{a_{j-1}c_j}{a_j}}, \quad j \geq 1, \tag{3.25}$$

with $\{a_j, b_j, c_j\}$ being the coefficients of the three-term recurrence relation (3.12), namely,

$$p_{j+1}(x) = (a_j x - b_j)p_j(x) - c_j p_{j-1}(x), \quad j \geq 0, \tag{3.26}$$

with $p_{-1} := 0$.

Proof. We first normalize the orthogonal polynomials $\{p_j\}$ by defining

$$\tilde{p}_j(x) = \frac{1}{\sqrt{\gamma_j}} p_j(x) \quad \text{with} \quad \gamma_j = \|p_j\|_\omega^2 \;\Rightarrow\; \|\tilde{p}_j\|_\omega = 1.$$

Thus, we have

$$\begin{aligned} x\tilde{p}_j &\overset{(3.26)}{=} \frac{c_j}{a_j}\sqrt{\frac{\gamma_{j-1}}{\gamma_j}}\tilde{p}_{j-1} + \frac{b_j}{a_j}\tilde{p}_j + \frac{1}{a_j}\sqrt{\frac{\gamma_{j+1}}{\gamma_j}}\tilde{p}_{j+1} \\ &\overset{(3.13)}{=} \frac{1}{a_{j-1}}\sqrt{\frac{\gamma_j}{\gamma_{j-1}}}\tilde{p}_{j-1} + \frac{b_j}{a_j}\tilde{p}_j + \frac{1}{a_j}\sqrt{\frac{\gamma_{j+1}}{\gamma_j}}\tilde{p}_{j+1} \\ &= \beta_j \tilde{p}_{j-1}(x) + \alpha_j \tilde{p}_j(x) + \beta_{j+1}\tilde{p}_{j+1}(x), \quad j \geq 0, \end{aligned} \tag{3.27}$$

where we denote

$$\alpha_j := \frac{b_j}{a_j}, \quad \beta_j := \frac{1}{a_{j-1}}\sqrt{\frac{\gamma_j}{\gamma_{j-1}}}.$$

By (3.13),

$$\frac{\gamma_j}{\gamma_{j-1}} = \frac{a_{j-1}c_j}{a_j} > 0 \;\Rightarrow\; \beta_j = \frac{1}{a_{j-1}}\sqrt{\frac{a_{j-1}c_j}{a_j}}.$$

Then we rewrite the recurrence relation (3.27) as

$$x\tilde{p}_j(x) = \beta_j\tilde{p}_{j-1}(x) + \alpha_j\tilde{p}_j(x) + \beta_{j+1}\tilde{p}_{j+1}(x), \quad j \ge 0.$$

We now take $j = 0, 1, \ldots, n$ to form a system with the matrix form

$$x\widetilde{\mathbf{P}}(x) = A_{n+1}\widetilde{\mathbf{P}}(x) + \beta_{n+1}\tilde{p}_{n+1}(x)\mathbf{E}_{n+1}, \tag{3.28}$$

where A_{n+1} is given by (3.24), and

$$\widetilde{\mathbf{P}}(x) = \big(\tilde{p}_0(x), \tilde{p}_1(x), \ldots, \tilde{p}_n(x)\big)^T, \quad \mathbf{E}_{n+1} = (0, 0, \ldots, 0, 1)^T.$$

Since $\tilde{p}_{n+1}(x_j) = 0,\ 0 \le j \le n$, the system (3.28) at $x = x_j$ becomes

$$x_j\widetilde{\mathbf{P}}(x_j) = A_{n+1}\widetilde{\mathbf{P}}(x_j), \quad 0 \le j \le n. \tag{3.29}$$

Hence, $\{x_j\}_{j=0}^n$ are eigenvalues of the symmetric tridiagonal matrix A_{n+1}. □

An alternative approach for finding zeros of orthogonal polynomials is to use an iterative procedure. More precisely, let x_j^0 be an initial approximation to the zero x_j of $p_{n+1}(x)$. Then, one can construct an iterative scheme in the general form:

$$\begin{cases} x_j^{k+1} = x_j^k + D(x_j^k), \;\; 0 \le j \le n, \;\; k \ge 0, \\ \text{given } \{x_j^0\}_{j=0}^n, \text{ and a termination rule.} \end{cases} \tag{3.30}$$

The deviation $D(\cdot)$ classifies different types of iterative schemes. For instance, the Newton method is of second-order with

$$D(x) = -\frac{p_{n+1}(x)}{p'_{n+1}(x)}, \tag{3.31}$$

while the Laguerre method is a third-order scheme with

$$D(x) = -\frac{p_{n+1}(x)}{p'_{n+1}(x)} - \frac{p_{n+1}(x)p''_{n+1}(x)}{2(p'_{n+1}(x))^2}. \tag{3.32}$$

Higher-order schemes can be constructed by using higher-order derivatives of p_{n+1}.

The success of an iterative method often depends on how good is the initial guess. If the initial approximation is not sufficiently close, the algorithm may converge to other unwanted values or even diverge. For a thorough discussion on how to find zeros of polynomials, we refer to Pan (1997) and the references therein.

3.1.4 Gauss-Type Quadratures

We now discuss the relations between orthogonal polynomials and Gauss-type integration formulas. The mechanism of a Gauss-type quadrature is to seek the best numerical approximation of an integral by selecting optimal nodes at which the integrand is evaluated. It belongs to the family of the numerical quadratures:

$$\int_a^b f(x)\omega(x)dx = \sum_{j=0}^{N} f(x_j)\omega_j + E_N[f], \tag{3.33}$$

where $\{x_j, \omega_j\}_{j=0}^N$ are the quadrature nodes and weights, and $E_N[f]$ is the quadrature error. If $E_N[f] \equiv 0$, we say the quadrature formula (3.33) is *exact* for f.

Hereafter, we assume that the nodes $\{x_j\}_{j=0}^N$ are distinct. If $f(x) \in C^{N+1}[a,b]$, we have (see, e.g., Davis and Rabinowitz (1984)):

$$E_N[f] = \frac{1}{(N+1)!}\int_a^b f^{(N+1)}(\xi(x))\prod_{i=0}^{N}(x-x_i)dx, \tag{3.34}$$

where $\xi(x) \in [a,b]$. The Lagrange basis polynomials associated with $\{x_j\}_{j=0}^N$ are given by

$$h_j(x) = \prod_{i=0;i\neq j}^{N} \frac{x-x_i}{x_j-x_i}, \quad 0 \le j \le N, \tag{3.35}$$

so taking $f(x) = h_j$ in (3.33) and using (3.34), we find the quadrature weights:

$$\omega_j = \int_a^b h_j(x)\omega(x)dx, \quad 0 \le j \le N. \tag{3.36}$$

We say that the integration formula (3.33)–(3.36) has a degree of precision (DOP) m, if there holds

$$E_N[p] = 0,\ \ \forall p \in P_m \ \text{ but } \exists q \in P_{m+1} \text{ such that } \ E_N[q] \neq 0. \tag{3.37}$$

In general, for any $N+1$ distinct nodes $\{x_j\}_{j=0}^N \subseteq (a,b)$, the DOP of (3.33)–(3.36) is between N and $2N+1$. Moreover, if the nodes $\{x_k\}_{k=0}^N$ are chosen as zeros of the polynomial p_{N+1} orthogonal with respect to ω, this rule enjoys the maximum degree of precision $2N+1$. Such a rule is known as the Gauss quadrature.

Theorem 3.5. (Gauss quadrature) *Let $\{x_j\}_{j=0}^N$ be the set of zeros of the orthogonal polynomial p_{N+1}. Then there exists a unique set of quadrature weights $\{\omega_j\}_{j=0}^N$, defined by* (3.36)*, such that*

$$\int_a^b p(x)\omega(x)dx = \sum_{j=0}^{N} p(x_j)\omega_j, \quad \forall p \in P_{2N+1}, \tag{3.38}$$

where the quadrature weights are all positive and given by

$$\omega_j = \frac{k_{N+1}}{k_N} \frac{\|p_N\|_\omega^2}{p_N(x_j) p'_{N+1}(x_j)}, \quad 0 \le j \le N, \tag{3.39}$$

where k_j is the leading coefficient of the polynomial p_j.

Proof. Let $\{h_j\}_{j=0}^N$ be the Lagrange basis polynomials defined in (3.35). It is clear that

$$P_N = \text{span}\{h_j : 0 \le j \le N\} \;\Rightarrow\; p(x) = \sum_{j=0}^N p(x_j) h_j(x), \;\; \forall p \in P_N.$$

Hence,

$$\int_a^b p(x)\omega(x)dx = \sum_{j=0}^N p(x_j) \int_a^b h_j(x)\omega(x)dx \overset{(3.36)}{=} \sum_{j=0}^N p(x_j)\omega_j, \tag{3.40}$$

which implies (3.38) is exact for any $p \in P_N$. In other words, the DOP of this rule is not less than N.

Next, for any $p \in P_{2N+1}$, we can write $p = rp_{N+1} + s$ where $r, s \in P_N$. In view of $p_{N+1}(x_j) = 0$, we have $p(x_j) = s(x_j)$ for $0 \le j \le N$. Since p_{N+1} is orthogonal to r (cf. Lemma 3.1) and $s \in P_N$, we find

$$\begin{aligned}\int_a^b p(x)\omega(x)dx &= \int_a^b s(x)\omega(x)dx \\ &= \sum_{j=0}^N s(x_j)\omega_j \overset{(3.40)}{=} \sum_{j=0}^N p(x_j)\omega_j, \quad \forall p \in P_{2N+1},\end{aligned} \tag{3.41}$$

which leads to (3.38).

Now, we show that $\omega_j > 0$ for $0 \le j \le N$. Taking $p(x) = h_j^2(x) \in P_{2N}$ in (3.41) leads to

$$0 < \int_a^b h_j^2(x)\omega(x)dx = \sum_{k=0}^N h_j^2(x_k)\omega_k = \omega_j, \quad 0 \le j \le N.$$

It remains to establish (3.39). Since $p_{N+1}(x_j) = 0$, taking $y = x_j$ and $n = N$ in the Christoff-Darboux formula (3.15) yields

$$p_N(x_j)\frac{p_{N+1}(x)}{x - x_j} = \frac{k_{N+1}}{k_N} \sum_{i=0}^N \frac{\|p_N\|_\omega^2}{\|p_i\|_\omega^2} p_i(x_j) p_i(x).$$

Multiplying the above formula by $\omega(x)$ and integrating the resulting identity over (a,b), we deduce from the orthogonality $(p_i, 1)_\omega = 0$, $i \ge 1$, that

$$\begin{aligned}&p_N(x_j)\int_a^b \frac{p_{N+1}(x)}{x - x_j}\omega(x)dx \\ &\quad = \frac{k_{N+1}}{k_N}\|p_N\|_\omega^2 \frac{(p_0(x_j), p_0)_\omega}{\|p_0\|_\omega^2} = \frac{k_{N+1}}{k_N}\|p_N\|_\omega^2.\end{aligned} \tag{3.42}$$

Note that the Lagrange basis polynomial $h_j(x)$ in (3.35) can be expressed as

$$h_j(x) = \frac{p_{N+1}(x)}{p'_{N+1}(x_j)(x-x_j)}. \tag{3.43}$$

Plugging it into (3.42) gives

$$\begin{aligned} p_N(x_j)\int_a^b \frac{p_{N+1}(x)}{x-x_j}\omega(x)dx &= p_N(x_j)p'_{N+1}(x_j)\int_a^b h_j(x)\omega(x)dx \\ &= p_N(x_j)p'_{N+1}(x_j)\omega_j = \frac{k_{N+1}}{k_N}\|p_N\|_\omega^2, \end{aligned} \tag{3.44}$$

which implies (3.39). □

The above fundamental theorem reveals that the optimal abscissas of the Gauss quadrature formula are precisely the zeros of the orthogonal polynomial for the same interval and weight function. The Gauss quadrature is optimal because it fits all polynomials up to degree $2N+1$ exactly, and it is impossible to find any set of $\{x_j,\ \omega_j\}_{j=0}^N$ such that (3.38) holds for all $p \in P_{2N+2}$ (see Problem 3.3).

With the exception of a few special cases, like the Chebyshev polynomials, no explicit expressions of the quadrature nodes and weights are available. Theorem 3.4 provides an efficient approach to compute the nodes $\{x_j\}_{j=0}^N$, through finding the eigenvalues of the symmetric tridiagonal matrix A_{N+1} defined in (3.24). The following result indicates that the weights $\{\omega_j\}_{j=0}^N$ can be computed from the first component of the eigenvectors of A_{N+1}.

Theorem 3.6. *Let*

$$\mathbf{Q}(x_j) = \big(Q_0(x_j), Q_1(x_j), \ldots, Q_N(x_j)\big)^T$$

be the orthonormal eigenvector of A_{N+1} corresponding to the eigenvalue x_j, i.e.,

$$A_{N+1}\mathbf{Q}(x_j) = x_j\mathbf{Q}(x_j) \quad \textit{with} \quad \mathbf{Q}(x_j)^T\mathbf{Q}(x_j) = 1.$$

Then the weights $\{\omega_j\}_{j=0}^N$ can be computed from the first component of the eigenvector $\mathbf{Q}(x_j)$ by using the formula:

$$\omega_j = \big[Q_0(x_j)\big]^2\int_a^b \omega(x)dx, \quad 0 \le j \le N. \tag{3.45}$$

Proof. Using the Christoffel-Darboux formula (3.16) and the fact that $p_{N+1}(x_j) = 0$, we derive from (3.39) the following alternative expression of the weights:

$$\begin{aligned} \omega_j^{-1} &\overset{(3.39)}{=} \frac{k_N}{k_{N+1}}\frac{p_N(x_j)p'_{N+1}(x_j)}{\|p_N\|_\omega^2} \overset{(3.16)}{=} \sum_{n=0}^N \frac{p_n^2(x_j)}{\|p_n\|_\omega^2} \\ &= \widetilde{\mathbf{P}}(x_j)^T\widetilde{\mathbf{P}}(x_j), \quad 0 \le j \le N, \end{aligned} \tag{3.46}$$

where

$$\widetilde{\mathbf{P}}(x_j) = \big(\tilde{p}_0(x_j), \tilde{p}_1(x_j), \ldots, \tilde{p}_N(x_j)\big)^T \text{ with } \tilde{p}_n = \frac{p_n}{\|p_n\|_\omega}.$$

The identity (3.46) can be rewritten as

$$\omega_j \widetilde{\mathbf{P}}(x_j)^T \widetilde{\mathbf{P}}(x_j) = 1, \quad 0 \le j \le N.$$

On the other hand, we deduce from (3.29) that $\widetilde{\mathbf{P}}(x_j)$ is an eigenvector corresponding to the eigenvalue x_j. Therefore,

$$\mathbf{Q}(x_j) = \sqrt{\omega_j}\, \widetilde{\mathbf{P}}(x_j), \quad 0 \le j \le N, \tag{3.47}$$

is the unit eigenvector corresponding to the eigenvalue x_j. Equating the first components (3.47) yields

$$\omega_j = \left[\frac{Q_0(x_j)}{\tilde{p}_0(x_j)}\right]^2 = \frac{\|p_0\|_\omega^2}{\big[p_0(x_j)\big]^2}\big[Q_0(x_j)\big]^2 = \big[Q_0(x_j)\big]^2 \int_a^b \omega(x)dx, \quad 0 \le j \le N.$$

This completes the proof. □

Notice that all the nodes of the Gauss formula lie in the interior of the interval (a,b). This makes it difficult to impose boundary conditions. Below, we consider the Gauss-Radau or Gauss-Lobatto quadratures which include either one or both endpoints as a node(s).

We start with the Gauss-Radau quadrature. Assuming we would like to include the left endpoint $x = a$ in the quadrature, we define

$$q_N(x) = \frac{p_{N+1}(x) + \alpha_N p_N(x)}{x-a} \text{ with } \alpha_N = -\frac{p_{N+1}(a)}{p_N(a)}. \tag{3.48}$$

It is obvious that $q_N \in P_N$, and for any $r_{N-1} \in P_{N-1}$, we derive from Lemma 3.1 that

$$\begin{aligned} &\int_a^b q_N(x) r_{N-1}(x)\omega(x)(x-a)dx \\ &\quad = \int_a^b \big(p_{N+1}(x) + \alpha_N p_N(x)\big)\, r_{N-1}(x)\omega(x)dx = 0. \end{aligned} \tag{3.49}$$

Hence, $\big\{q_N : N \ge 0\big\}$ defines a sequence of polynomials orthogonal with respect to the weight function $\tilde{\omega}(x) := \omega(x)(x-a)$, and the leading coefficient of q_N is k_{N+1}.

Theorem 3.7. (Gauss-Radau quadrature) *Let $x_0 = a$ and $\{x_j\}_{j=1}^N$ be the zeros of q_N defined in (3.48). Then there exists a unique set of quadrature weights $\{\omega_j\}_{j=0}^N$, defined by (3.36), such that*

$$\int_a^b p(x)\omega(x)dx = \sum_{j=0}^N p(x_j)\omega_j, \quad \forall p \in P_{2N}. \tag{3.50}$$

Moreover, the quadrature weights are all positive and can be expressed as

$$\omega_0 = \frac{1}{q_N(a)} \int_a^b q_N(x)\omega(x)dx, \tag{3.51a}$$

$$\omega_j = \frac{1}{x_j - a} \frac{k_{N+1}}{k_N} \frac{\|q_{N-1}\|^2_{\tilde{\omega}}}{q_{N-1}(x_j)q'_N(x_j)}, \quad 1 \le j \le N. \tag{3.51b}$$

Proof. The proof is similar to that of Theorem 3.5, so we shall only sketch it below.
Obviously, for any $p \in P_N$,

$$\int_a^b p(x)\omega(x)dx = \sum_{j=0}^{N} p(x_j) \int_a^b h_j(x)\omega(x)dx \overset{(3.36)}{=} \sum_{j=0}^{N} p(x_j)\omega_j. \tag{3.52}$$

Hence, the DOP is at least N.
Next, for any $p \in P_{2N}$, we write

$$p = (x-a)rq_N + s, \quad r \in P_{N-1},\ s \in P_N.$$

Since $(x-a)q_N(x)\big|_{x=x_j} = 0$, we have $p(x_j) = s(x_j)$ for $0 \le j \le N$. Therefore, we deduce from (3.49) that

$$\begin{aligned}\int_a^b p(x)\omega(x)dx &= \int_a^b s(x)\omega(x)dx \\ &= \sum_{j=0}^{N} s(x_j)\omega_j = \sum_{j=0}^{N} p(x_j)\omega_j, \quad \forall p \in P_{2N}.\end{aligned}$$

Taking $p(x) = h_k^2(x) \in P_{2N}$ in the above identities, we conclude that $\omega_k > 0$ for $0 \le k \le N$.
Note that the Lagrange basis polynomials take the form

$$\begin{aligned}h_j(x) &= \frac{(x-a)q_N(x)}{\big((x-a)q_N(x)\big)'\big|_{x=x_j}(x-x_j)} \\ &= \frac{(x-a)q_N(x)}{\big(q_N(x_j) + (x_j-a)q'_N(x_j)\big)(x-x_j)}, \quad 0 \le j \le N.\end{aligned}$$

Hence, letting $j = 0$, we derive (3.51a) from the definition of ω_0, and for $1 \le j \le N$,

$$\omega_j = \int_a^b h_j(x)\omega(x)dx = \frac{1}{x_j - a} \int_a^b \frac{q_N(x)}{q'_N(x_j)(x-x_j)} \tilde{\omega}(x)dx.$$

Recall that $\{q_n\}$ are orthogonal with respect to $\tilde{\omega}$, so the integral part turns out to be the weight of the Gauss quadrature associated with N nodes being the zeros of $q_N(x)$. Hence, (3.51b) follows from the formula (3.39). □

Remark 3.1. *Similarly, a second Gauss-Radau quadrature can be constructed if we want to include the right endpoint $x = b$ instead of the left endpoint $x = a$.*

We now turn to the Gauss-Lobatto quadrature, whose nodes include two endpoints $x = a, b$. In this case, we choose α_N and β_N such that

$$p_{N+1}(x) + \alpha_N p_N(x) + \beta_N p_{N-1}(x) = 0 \text{ for } x = a, b, \tag{3.53}$$

and set

$$z_{N-1}(x) = \frac{p_{N+1}(x) + \alpha_N p_N(x) + \beta_N p_{N-1}(x)}{(x-a)(b-x)}. \tag{3.54}$$

It is clear that $z_{N-1} \in P_{N-1}$ and for any $r_{N-2} \in P_{N-2}$, we derive from Lemma 3.1 that

$$\begin{aligned} &\int_a^b z_{N-1} r_{N-2}(x-a)(b-x)\omega\, dx \\ &= \int_a^b \left(p_{N+1} + \alpha_N p_N + \beta_N p_{N-1}\right) r_{N-2}\,\omega\, dx = 0. \end{aligned} \tag{3.55}$$

Hence, $\left\{z_{N-1} : N \geq 1\right\}$ defines a sequence of polynomials orthogonal with respect to the weight function $\hat{\omega}(x) := (x-a)(b-x)\omega(x)$, and the leading coefficient of z_{N-1} is $-k_{N+1}$.

Theorem 3.8. (Gauss-Lobatto quadrature) *Let $x_0 = a$, $x_N = b$ and $\{x_j\}_{j=1}^{N-1}$ be the zeros of z_{N-1} in (3.53)–(3.54). Then there exists a unique set of quadrature weights $\{\omega_j\}_{j=0}^{N}$, defined by (3.36), such that*

$$\int_a^b p(x)\omega(x)dx = \sum_{j=0}^{N} p(x_j)\omega_j, \quad \forall p \in P_{2N-1}, \tag{3.56}$$

where the quadrature weights are expressed as

$$\omega_0 = \frac{1}{(b-a)z_{N-1}(a)} \int_a^b (b-x) z_{N-1}(x)\omega(x)dx, \tag{3.57a}$$

$$\omega_j = \frac{1}{(x_j-a)(b-x_j)} \frac{k_{N+1}}{k_N} \frac{\|z_{N-2}\|_{\hat{\omega}}^2}{z_{N-2}(x_j) z'_{N-1}(x_j)}, \quad 1 \leq j \leq N-1, \tag{3.57b}$$

$$\omega_N = \frac{1}{(b-a)z_{N-1}(b)} \int_a^b (x-a) z_{N-1}(x)\omega(x)dx. \tag{3.57c}$$

Moreover, we have $\omega_j > 0$ for $1 \leq j \leq N-1$.

Proof. The exactness (3.56) and the formulas of the weights can be derived in a similar fashion as in Theorem 3.7, so we skip the details. Here, we just verify $\omega_j > 0$ for $1 \leq j \leq N-1$ by using a different approach. Since $\{z_{N-1}\}$ are orthogonal with respect to the weight function $\hat{\omega}$, and $z_{N-1}(x_j) = 0$ for $1 \le j \le N-1$, we obtain from the Christoff-Darboux formula (3.16) that

$$\frac{k_N}{k_{N+1}} z_{N-2}(x_j) z'_{N-1}(x_j) = \sum_{j=0}^{N-2} \frac{\|z_{N-2}\|^2_{\hat{\omega}}}{\|z_j\|^2_{\hat{\omega}}} z_j^2(x_j) > 0, \;\; 1 \le j \le N-1.$$

Inserting it into the formula (3.57b) leads to $\omega_j > 0$ for $1 \le j \le N-1$. □

The Gauss-type quadrature formulas provide powerful tools for evaluating integrals and inner products in a spectral method. They also play an important role in spectral differentiations as to be shown later.

3.1.5 Interpolation and Discrete Transforms

Let $\{x_j, \omega_j\}_{j=0}^N$ be a set of Gauss, Gauss-Radau or Gauss-Lobatto quadrature nodes and weights. We define the corresponding discrete inner product and norm as

$$\langle u, v \rangle_{N,\omega} := \sum_{j=0}^{N} u(x_j) v(x_j) \omega_j, \quad \|u\|_{N,\omega} := \sqrt{\langle u, u \rangle_{N,\omega}}. \tag{3.58}$$

Note that $\langle \cdot,\, \cdot \rangle_{N,\omega}$ is an approximation to the continuous inner product $(\cdot, \cdot)_\omega$, and the exactness of Gauss-type quadrature formulas implies

$$\langle u, v \rangle_{N,\omega} = (u, v)_\omega, \quad \forall u \cdot v \in P_{2N+\delta}, \tag{3.59}$$

where $\delta = 1$, 0 and -1 for the Gauss, Gauss-Radau and Gauss-Lobatto quadrature, respectively.

Definition 3.1. *For any $u \in C(\Lambda)$, we define the interpolation operator $I_N : C(\Lambda) \to P_N$ such that*

$$(I_N u)(x_j) = u(x_j), \qquad 0 \le j \le N, \tag{3.60}$$

where $\Lambda = (a,b)$, $[a,b)$, $[a,b]$ for the Gauss, Gauss-Radau and Gauss-Lobatto quadrature, respectively.

The interpolation condition (3.60) implies that $I_N p = p$ for all $p \in P_N$. On the other hand, since $I_N u \in P_N$, we can write

$$(I_N u)(x) = \sum_{n=0}^{N} \tilde{u}_n p_n(x), \tag{3.61}$$

which is the counterpart of the discrete Fourier series (2.20) and may be referred to as the *discrete polynomial series*. By taking the discrete inner product of (3.61) with $\{p_k\}_{k=0}^N$, we can determine the coefficients $\{\tilde{u}_n\}$ by using (3.60) and (3.59). More precisely, we have

Theorem 3.9.

$$\tilde{u}_n = \frac{1}{\gamma_n} \sum_{j=0}^{N} u(x_j) p_n(x_j) \omega_j, \quad 0 \le n \le N, \tag{3.62}$$

where $\gamma_n = \|p_n\|_\omega^2$ *for* $0 \le n \le N-1$, *and*

$$\gamma_N = \begin{cases} \|p_N\|_\omega^2, & \textit{for Gauss and Gauss-Radau,} \\ \langle p_N, p_N \rangle_{N,\omega}, & \textit{for Gauss-Lobatto.} \end{cases} \tag{3.63}$$

The formula (3.62)-(3.63) defines the *forward discrete polynomial transform* as in the Fourier case, which transforms the physical values $\{u(x_j)\}_{j=0}^N$ to the expansion coefficients $\{\tilde{u}_n\}_{n=0}^N$. Conversely, the *backward* (*or inverse*) *discrete polynomial transform* is formulated by

$$u(x_j) = (I_N u)(x_j) = \sum_{n=0}^{N} \tilde{u}_n p_n(x_j), \quad 0 \le j \le N, \tag{3.64}$$

which takes the expansion coefficients $\{\tilde{u}_n\}_{n=0}^N$ to the physical values $\{u(x_j)\}_{j=0}^N$.

We see that if the matrices $\big(p_n(x_j)\big)_{0\le n,j\le N}$ and/or $\big(\gamma_n^{-1} p_n(x_j)\omega_j\big)_{0\le n,j\le N}$ are precomputed, then the discrete transforms (3.62) and (3.64) can be manipulated directly by a standard matrix–vector multiplication routine in about N^2 flops. Since discrete transforms are frequently used in spectral codes, it is desirable to reduce the computational complexity, especially for multidimensional cases. In particular, the Fast Fourier Transform (FFT) (cf. Cooley and Tukey (1965)) and discrete Chebyshev transform (treated as a Fourier-cosine transform) can be accomplished by $O(N \log_2 N)$ operations. However, with the advent of more powerful computers, this aspect should not be a big concern for moderate scale problems.

3.1.6 Differentiation in the Physical Space

Now, we are ready to address an important issue – polynomial-based spectral differentiation techniques. As with the Fourier cases, they can be performed in either the physical space or the frequency space.

Let us start with the implementation in the physical space. Assume that $u \in P_N$ is an approximation of the unknown solution U. Let $\{h_j\}_{j=0}^N$ be the Lagrange basis polynomials associated with a set of Gauss-type points $\{x_j\}_{j=0}^N$. Clearly,

$$u(x) = \sum_{j=0}^{N} u(x_j) h_j(x). \tag{3.65}$$

Hence, differentiating it m times leads to

$$u^{(m)}(x_k) = \sum_{j=0}^{N} h_j^{(m)}(x_k) u(x_j), \quad 0 \le k \le N. \tag{3.66}$$

Let us denote

$$\mathbf{u}^{(m)} := \big(u^{(m)}(x_0), u^{(m)}(x_1), \ldots, u^{(m)}(x_N)\big)^T, \quad \mathbf{u} := \mathbf{u}^{(0)};$$
$$D^{(m)} := \Big(d_{kj}^{(m)}\Big)_{0\le k,j\le N}, \quad D := D^{(1)}. \tag{3.67}$$

Different from the Fourier case, the higher-order differentiation matrix in this context can be computed by a product of the first-order one.

Theorem 3.10.

$$D^{(m)} = DD\ldots D = D^m, \quad m \ge 1, \tag{3.68}$$

and

$$\mathbf{u}^{(m)} = D^m \mathbf{u}, \quad m \ge 1. \tag{3.69}$$

Proof. Differentiating (3.65) gives

$$u'(x) = \sum_{l=0}^{N} u(x_l) h_l'(x). \tag{3.70}$$

Taking $u = h_j' \in P_{N-1}$ in the above equation leads to

$$h_j''(x) = \sum_{l=0}^{N} h_l'(x) h_j'(x_l).$$

Hence,

$$d_{kj}^{(2)} = h_j''(x_k) = \sum_{l=0}^{N} h_l'(x_k) h_j'(x_l) = \sum_{l=0}^{N} d_{kl}^{(1)} d_{lj}^{(1)},$$

which implies

$$D^{(2)} = DD = D^2. \tag{3.71}$$

Similarly, taking $u = h_j^{(i)}$ in (3.70) leads to

$$d_{kj}^{(i+1)} = h_j^{(i+1)}(x_k) = \sum_{l=0}^{N} h_l'(x_k) h_j^{(i)}(x_l) = \sum_{l=0}^{N} d_{kl}^{(1)} d_{lj}^{(i)}.$$

Therefore,

$$D^{(i+1)} = DD^{(i)}, \quad i \ge 1, \tag{3.72}$$

which yields (3.68).

Finally, (3.69) can be written in matrix form as in (3.66). □

Thanks to Theorem 3.10, it suffices to compute the first-order differentiation matrix D. We present below the explicit formulas for the entries of D.

Theorem 3.11. *The entries of D are determined by*

$$d_{kj} = h_j'(x_k) = \begin{cases} \dfrac{Q'(x_k)}{Q'(x_j)} \dfrac{1}{x_k - x_j}, & \text{if } k \neq j, \\ \dfrac{Q''(x_k)}{2Q'(x_k)}, & \text{if } k = j, \end{cases} \tag{3.73}$$

where

$$Q(x) = p_{N+1}(x),\ (x-a)q_N(x),\ (x-a)(b-x)z_{N-1}(x) \tag{3.74}$$

are the quadrature polynomials (cf. (3.48) and (3.54)) of the Gauss, Gauss-Radau and Gauss-Lobatto quadrature, respectively.

Proof. The Lagrange basis polynomials can be expressed as

$$h_j(x) = \frac{Q(x)}{Q'(x_j)(x - x_j)}, \quad 0 \le j \le N. \tag{3.75}$$

Differentiating (3.75) and using the fact $Q(x_j) = 0$ lead to

$$d_{kj} = h_j'(x_k) = \frac{Q'(x_k)}{Q'(x_j)} \frac{1}{x_k - x_j}, \quad \forall k \neq j.$$

Applying the L'Hopital's rule twice yields

$$d_{kk} = \lim_{x \to x_k} h_k'(x) = \frac{1}{Q'(x_k)} \lim_{x \to x_k} \frac{Q'(x)(x - x_k) - Q(x)}{(x - x_k)^2} = \frac{Q''(x_k)}{2Q'(x_k)}.$$

This completes the proof. □

Therefore, having precomputed the first-order differentiation matrix, the differentiation in the physical space can be carried out through matrix–matrix and matrix–vector multiplications.

3.1.7 Differentiation in the Frequency Space

Differentiation in the frequency space is to express the expansion coefficients of the derivatives of a function in terms of expansion coefficients of the function itself. More precisely, given $u \in P_N$, instead of using the Lagrange basis polynomials, we expand u in terms of the orthogonal polynomials:

$$u(x) = \sum_{n=0}^{N} \hat{u}_n p_n(x). \tag{3.76}$$

Using the orthogonality, we can determine the coefficients by

$$\hat{u}_n = \frac{1}{\|p_n\|_\omega^2} \int_a^b u(x) p_n(x) \omega(x) dx, \quad 0 \le n \le N. \tag{3.77}$$

Since $u' \in P_{N-1}$, we have

$$u'(x) = \sum_{n=0}^{N} \hat{u}_n^{(1)} p_n(x) \quad \text{with} \quad \hat{u}_N^{(1)} = 0. \tag{3.78}$$

In order to express $\{\hat{u}_n^{(1)}\}_{n=0}^N$ in terms of $\{\hat{u}_n\}_{n=0}^N$, we assume that $\{p_n'\}$ are also orthogonal. Indeed, this property holds for the classical orthogonal polynomials such as the Legendre, Chebyshev, Jacobi, Laguerre and Hermite polynomials. In other words, $\{p_n'\}$ satisfy the three-term recurrence relation due to Corollary 3.1:

$$p_{n+1}'(x) = \left(a_n^{(1)} x - b_n^{(1)}\right) p_n'(x) - c_n^{(1)} p_{n-1}'(x). \tag{3.79}$$

Differentiating the three-term recurrence relation (3.12) and using (3.79), we derive

$$p_n(x) = \tilde{a}_n p_{n-1}'(x) + \tilde{b}_n p_n'(x) + \tilde{c}_n p_{n+1}'(x). \tag{3.80}$$

The coefficients $\left\{\hat{u}_n^{(1)}\right\}$ in (3.78) can be computed by the following backward recurrence formulas.

Theorem 3.12.

$$\begin{aligned} \hat{u}_{n-1}^{(1)} &= \frac{1}{\tilde{c}_{n-1}} \left[\hat{u}_n - \tilde{b}_n \hat{u}_n^{(1)} - \tilde{a}_{n+1} \hat{u}_{n+1}^{(1)}\right], \quad n = N-1, \ldots, 1, \\ \hat{u}_N^{(1)} &= 0, \quad \hat{u}_{N-1}^{(1)} = \frac{1}{\tilde{c}_{N-1}} \hat{u}_N. \end{aligned} \tag{3.81}$$

Proof. By (3.78) and (3.80),

$$\begin{aligned} u' &= \sum_{n=0}^{N-1} \hat{u}_n^{(1)} p_n = \sum_{n=0}^{N-1} \hat{u}_n^{(1)} [\tilde{a}_n p_{n-1}' + \tilde{b}_n p_n' + \tilde{c}_n p_{n+1}'] \\ &= \sum_{n=1}^{N-1} [\tilde{c}_{n-1} \hat{u}_{n-1}^{(1)} + \tilde{b}_n \hat{u}_n^{(1)} + \tilde{a}_{n+1} \hat{u}_{n+1}^{(1)}] p_n' + \tilde{c}_{N-1} \hat{u}_{N-1}^{(1)} p_N'. \end{aligned}$$

On the other hand, by (3.76),

$$u'(x) = \sum_{n=1}^{N} \hat{u}_n p_n'(x).$$

By the (assumed) orthogonality of $\{p_n'\}$, we are able to equate the coefficients of p_n' in the above two expressions, which leads to (3.81). □

Higher-order differentiations in the frequency space can be carried out by using the formula (3.81) repeatedly. It is important to point out that *spectral differentiations* together with *discrete transforms* form the basic ingredients for the so-called "*pseudo-spectral technique*" (particularly useful for nonlinear problems): *the differentiations* are manipulated in the frequency space, the inner products are computed in the physical space, and both spaces are communicated through discrete transforms.

3.1.8 Approximability of Orthogonal Polynomials

We now briefly review some general polynomial approximation results. One can find their proofs from standard books on approximation theory (see, for instance, Timan (1994), Cheney (1998)).

The first fundamental result is the remarkable *Weierstrass Theorem*, which states that any continuous function in a finite interval can be uniformly approximated by an algebraic polynomial.

Theorem 3.13. *Let (a,b) be a finite interval. Then for any $u \in C[a,b]$, and any $\varepsilon > 0$, there exist $n \in \mathbb{N}$ and $p_n \in P_n$ such that*

$$\|u - p_n\|_{L^\infty(a,b)} < \varepsilon. \tag{3.82}$$

This theorem forms the cornerstone of the classical polynomial approximation theory. The construction of p_n essentially relies on the solution of the best approximation problem:

$$\begin{cases} \text{Given a fixed } n \in \mathbb{N}, \text{ find } p_n^* \in P_n, \text{ such that} \\ \|u - p_n^*\|_{L^\infty(a,b)} = \inf_{p_n \in P_n} \|u - p_n\|_{L^\infty(a,b)}. \end{cases} \tag{3.83}$$

This problem admits a unique solution, and as a consequence of Theorem 3.13, p_n^* uniformly converges to u as $n \to \infty$. However, the derivation of the best uniform approximation polynomial p_n^* is nontrivial, since a strong uniform norm is involved in (3.83), whereas the best approximation problem in the L^2-sense is easier to solve.

Theorem 3.14. *Let $I = (a,b)$ be a finite or an infinite interval. Then for any $u \in L^2_\omega(I)$ and $n \in \mathbb{N}$, there exists a unique $q_n^* \in P_n$, such that*

$$\|u - q_n^*\|_\omega = \inf_{q_n \in P_n} \|u - q_n\|_\omega, \tag{3.84}$$

where

$$q_n^*(x) = \sum_{k=0}^{n} \hat{u}_k p_k(x) \quad \text{with} \quad \hat{u}_k = \frac{(u, p_k)_\omega}{\|p_k\|_\omega^2}, \tag{3.85}$$

and $\{p_k\}_{k=0}^n$ *forms an* L^2_ω*-orthogonal basis of* P_n.

In particular, we denote the best approximation polynomial q_n^* by $\pi_n u$, which is the L^2_ω-orthogonal projection of u, and is characterized by the projection theorem

$$\|u - \pi_n u\|_\omega = \inf_{q_n \in P_n} \|u - q_n\|_\omega. \tag{3.86}$$

Equivalently, the L^2_ω-orthogonal projection can be defined by

$$(u - \pi_n u, \phi)_\omega = 0, \quad \forall \phi \in P_n, \tag{3.87}$$

so $\pi_n u$ is the first $n+1$-term truncation of the series $u = \sum_{k=0}^{\infty} \hat{u}_k p_k(x)$.

It is interesting to notice that a result similar to the Weierstrass theorem holds on infinite intervals, if suitable conditions are imposed on the growth of the given function u (cf. Funaro (1992)).

Theorem 3.15. *If* $u \in C[0,\infty)$ *and for certain* $\delta > 0$, u *satisfies*

$$u(x)e^{-\delta x} \to 0, \quad \text{as } x \to \infty,$$

then for any $\varepsilon > 0$, *there exist an* $n \in \mathbb{N}$ *and* $p_n \in P_n$ *such that*

$$|u(x) - p_n(x)|e^{-\delta x} \le \varepsilon, \quad \forall x \in [0,\infty).$$

Similar result holds on $(-\infty,\infty)$, *if we replace* $e^{-\delta x}$ *by* $e^{-\delta x^2}$.

3.1.8.1 A Short Summary of this Section

We presented some basic knowledge of orthogonal polynomials, which is mostly relevant to spectral approximations. We also set up a general framework for the study of each specific family of orthogonal polynomials to be presented in the forthcoming sections as tabulated in Table 3.1.

Table 3.1 List of orthogonal polynomials

	Symbol	Interval	Weight function	Section
Jacobi	$J_n^{\alpha,\beta}$	$(-1,1)$	$(1-x)^\alpha(1+x)^\beta, \alpha,\beta > -1$	3.2
Legendre	L_n	$(-1,1)$	1	3.3
Chebyshev	T_n	$(-1,1)$	$1/\sqrt{1-x^2}$	3.4
Laguerre	$\mathscr{L}_n^{(\alpha)}$	$(0,+\infty)$	$x^\alpha e^{-x}, \alpha > -1$	7.1
Hermite	H_n	$(-\infty,+\infty)$	e^{-x^2}	7.2