2
Best Approximation

2.1 The Legendre Polynomials

Given a function f defined on $[-1, 1]$, let us write

$$\|f\| = \left(\int_{-1}^{1} [f(x)]^2 dx\right)^{1/2}. \tag{2.1}$$

We call $\|f\|$ the *square norm* of f. It can be thought of as a measure of the "size" of f. The reason for taking the square root in (2.1) is so that the norm satisfies the condition

$$\|\lambda f\| = |\lambda| \cdot \|f\|, \tag{2.2}$$

for all real λ. The square norm, which is analogous to the notion of length in n-dimensional Euclidean space, obviously satisfies the *positivity* condition

$$\|f\| > 0 \quad \text{unless } f(x) \equiv 0, \text{ the zero function, when } \|f\| = 0, \tag{2.3}$$

and, not so obviously, satisfies the *triangle inequality*

$$\|f + g\| \leq \|f\| + \|g\|, \tag{2.4}$$

for all f and g. It is very easy to check that properties (2.2) and (2.3) hold for the square norm (2.1). The third property (2.4) is a little more difficult to justify. It may be verified by expressing the integrals as limits of sums and then applying the result in Problem 2.1.1. The square norm is a special case of a general norm, which we now define.

Definition 2.1.1 A *norm* $\|\cdot\|$ on a given *linear space* S is a mapping from S to the real numbers that satisfies the three properties given by (2.2), (2.3), and (2.4). ■

Note that a linear space contains a zero element, and the norm of the zero element is zero. Two examples of linear spaces are the linear space of n-dimensional vectors, and the linear space of continuous functions defined on a finite interval, say $[-1, 1]$, which we denote by $C[-1, 1]$. In the latter case the three best known norms are the square norm, defined by (2.1), the *maximum norm*, defined by

$$\|f\| = \max_{-1 \leq x \leq 1} |f(x)|, \tag{2.5}$$

and the norm defined by

$$\|f\| = \int_{-1}^{1} |f(x)| dx. \tag{2.6}$$

The three properties (2.2), (2.3), and (2.4), called the *norm axioms*, are all easily verified for the norms defined by (2.5) and (2.6). The norms given by (2.1) and (2.6) are special cases of the p-norm, defined by

$$\|f\| = \left(\int_{-1}^{1} |f(x)|^p dx \right)^{1/p}, \tag{2.7}$$

for any $p \geq 1$, and the maximum norm (2.5) is obtained by letting $p \to \infty$ in (2.7). The restriction $p \geq 1$ is necessary so that the p-norm satisfies the triangle inequality, which follows by expressing the integrals as limits of sums and applying Minkowski's inequality,

$$\left(\sum_{j=1}^{n} |x_j + y_j|^p \right)^{1/p} \leq \left(\sum_{j=1}^{n} |x_j|^p \right)^{1/p} + \left(\sum_{j=1}^{n} |y_j|^p \right)^{1/p}, \tag{2.8}$$

for $p \geq 1$. (See Davis [10] for a proof of (2.8).)

Example 2.1.1 Consider the linear space whose elements are the row vectors

$$\mathbf{x} = (x_1, x_2, \ldots, x_n),$$

where the x_j are all real, with the usual addition of vectors

$$\mathbf{x} + \mathbf{y} = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n).$$

We also have multiplication by a scalar, defined by

$$\lambda\mathbf{x} = (\lambda x_1, \lambda x_2, \ldots, \lambda x_n),$$

where λ is any real number. Then

$$\|\mathbf{x}\| = \max_{1 \le j \le n} |x_j|, \tag{2.9}$$

$$\|\mathbf{x}\| = |x_1| + \cdots + |x_n|, \tag{2.10}$$

$$\|\mathbf{x}\| = \left(x_1^2 + \cdots + x_n^2\right)^{1/2} \tag{2.11}$$

are all norms for this linear space. Except for the verification of the triangle inequality for the last norm (see Problem 2.1.1), it is easy to check that (2.9), (2.10), and (2.11) are indeed norms. (Corresponding to the zero function in (2.3) we have the zero vector, whose elements are all zero.) These vector norms are analogous to the norms (2.5), (2.6), and (2.1), respectively, given above for the linear space of functions defined on $[-1, 1]$. There is also the p-norm,

$$\|\mathbf{x}\| = \left(|x_1|^p + |x_2|^p + \cdots + |x_n|^p\right)^{1/p}, \tag{2.12}$$

for any $p \ge 1$, which is analogous to the norm defined by (2.7). It is easy to verify that the norm defined by (2.12) satisfies the properties (2.2) and (2.3), and we can apply Minkowski's inequality (2.8) to justify that it also satisfies the triangle inequality (2.4). If we put $p = 1$ and 2 in (2.12), we recover the norms (2.10) and (2.11), and if we let $p \to \infty$ in (2.12), we recover the norm (2.9). ■

Now, given f defined on $[-1, 1]$, let us seek the minimum value of $\|f - p\|$, for all $p \in P_n$, where $\|\cdot\|$ denotes the square norm. For *any* given norm, a $p \in P_n$ that minimizes $\|f - p\|$ is called a *best approximation* for f with respect to that norm. It can be shown (see Davis [10]) that for any norm and any $n \ge 0$, a best approximation always exists. Let us write

$$p(x) = \sum_{r=0}^{n} a_r q_r(x), \tag{2.13}$$

where $\{q_0, q_1, \ldots, q_n\}$ is some basis for P_n, so that any polynomial in P_n can be written as a sum of multiples of the q_r, as in (2.13). To find a best approximation, we can dispense with the square root in (2.1), since the problem of minimizing $\|f - p\|$ is equivalent to finding the minimum value of

$$\|f - p\|^2 = \int_{-1}^{1} [f(x) - p(x)]^2 dx = E(a_0, \ldots, a_n),$$

say, where p is given by (2.13). Thus we need to equate to zero the partial derivatives of E with respect to each a_s, and we obtain

$$0 = \frac{\partial}{\partial a_s} E(a_0, \ldots, a_n) = \int_{-1}^{1} 2[f(x) - p(x)] \cdot [-q_s(x)] dx,$$

for $0 \leq s \leq n$. This gives a system of linear equations to determine the coefficients a_s, and we will write these as

$$\sum_{r=0}^{n} c_{r,s} a_r = b_s, \quad 0 \leq s \leq n,$$

where

$$b_s = \int_{-1}^{1} f(x) q_s(x) dx, \quad 0 \leq s \leq n,$$

and

$$c_{r,s} = \int_{-1}^{1} q_r(x) q_s(x) dx, \quad 0 \leq r, s \leq n.$$

If we can now choose the basis $\{q_0, q_1, \ldots, q_n\}$ so that

$$\int_{-1}^{1} q_r(x) q_s(x) dx = 0, \;\; r \neq s, \;\; 0 \leq r, s \leq n, \tag{2.14}$$

then the above linear system will be immediately solved, giving

$$a_s = \int_{-1}^{1} f(x) q_s(x) dx \Big/ \int_{-1}^{1} [q_s(x)]^2 dx, \quad 0 \leq s \leq n. \tag{2.15}$$

Definition 2.1.2 The set of functions $\{q_0, q_1, \ldots, q_n\}$ is called an *orthogonal* basis if (2.14) holds, and if, in addition,

$$\int_{-1}^{1} [q_r(x)]^2 dx = 1, \quad 0 \leq r \leq n, \tag{2.16}$$

it is called an *orthonormal* basis. An orthogonal basis can obviously be made orthonormal by scaling each polynomial q_r appropriately. ■

We can construct the elements of an orthogonal basis, beginning with $q_0(x) = 1$ and choosing each q_k so that it satisfies the k conditions

$$\int_{-1}^{1} x^r q_k(x) dx = 0, \quad 0 \leq r < k. \tag{2.17}$$

We then say that q_k is orthogonal on $[-1, 1]$ to all polynomials in P_{k-1}, and thus (2.14) holds. To determine each $q_k \in P_k$ uniquely, we will scale q_k so that its coefficient of x^k is unity. We say that q_r and q_s are mutually orthogonal if $r \neq s$. These orthogonal polynomials are named after A. M. Legendre (1752–1833). Let us now write

$$q_k(x) = x^k + d_{k-1} x^{k-1} + \cdots + d_1 x + d_0,$$

and solve a system of k linear equations, derived from (2.17), to obtain the coefficients d_0, $d_1, \ldots, d_{k-1}$. Beginning with $q_0(x) = 1$, we find that

$q_1(x) = x$, and $q_2(x) = x^2 - \frac{1}{3}$. At this stage, we will refer to any multiples of the polynomials q_k as Legendre polynomials, although we will restrict the use of this name later to the particular *multiples* of these polynomials that assume the value 1 at $x = 1$. The following theorem shows that the Legendre polynomials satisfy a simple recurrence relation, and in Theorem 2.1.2 we will show that the recurrence relation can be simplified still further.

Theorem 2.1.1 The Legendre polynomials, scaled so that their leading coefficients are unity, satisfy the recurrence relation

$$q_{n+1}(x) = (x - \alpha_n)q_n(x) - \beta_n q_{n-1}(x), \tag{2.18}$$

where

$$\alpha_n = \int_{-1}^{1} x[q_n(x)]^2 dx \Big/ \int_{-1}^{1} [q_n(x)]^2 dx, \tag{2.19}$$

and

$$\beta_n = \int_{-1}^{1} [q_n(x)]^2 dx \Big/ \int_{-1}^{1} [q_{n-1}(x)]^2 dx, \tag{2.20}$$

for all $n \geq 1$, where $q_0(x) = 1$ and $q_1(x) = x$.

Proof. It is clear that q_0 and q_1 are mutually orthogonal. To complete the proof it will suffice to show that for $n \geq 1$, if $q_0, q_1, \ldots, q_n$ denote the Legendre polynomials of degree up to n, each with leading coefficient unity, then the polynomial q_{n+1} defined by (2.18), with α_n and β_n defined by (2.19) and (2.20), respectively, is orthogonal to all polynomials in P_n. Now, since q_{n-1} and q_n are orthogonal to all polynomials in P_{n-2} and P_{n-1}, respectively, it follows from the recurrence relation (2.18) that q_{n+1} is orthogonal to all polynomials in P_{n-2}. For if we multiply (2.18) throughout by $q_m(x)$ and integrate over $[-1, 1]$, it is clear from the orthogonality property that

$$\int_{-1}^{1} q_{n+1}(x)q_m(x)dx = \int_{-1}^{1} xq_n(x)q_m(x)dx = 0,$$

for $0 \leq m \leq n-2$, since we can write

$$xq_m(x) = q_{m+1}(x) + r_m(x),$$

where $r_m \in P_m$. The proof will be completed if we can show that q_{n+1} is orthogonal to q_{n-1} and q_n. If we multiply (2.18) throughout by $q_{n-1}(x)$, integrate over $[-1, 1]$, and use the fact that q_{n-1} and q_n are orthogonal, we obtain

$$\int_{-1}^{1} q_{n+1}(x)q_{n-1}(x)dx = \int_{-1}^{1} xq_n(x)q_{n-1}(x)dx - \beta_n \int_{-1}^{1} [q_{n-1}(x)]^2 dx.$$

Since $xq_{n-1}(x) = q_n(x) + r_{n-1}(x)$, where $r_{n-1} \in P_{n-1}$, it follows from the orthogonality property that

$$\int_{-1}^{1} xq_n(x)q_{n-1}(x)dx = \int_{-1}^{1} [q_n(x)]^2 dx.$$

From this and the above definition of β_n, it follows that q_{n+1} is orthogonal to q_{n-1}. If we now multiply (2.18) throughout by $q_n(x)$ and integrate over $[-1, 1]$, it then follows from the orthogonality of q_{n-1} and q_n, and the definition of α_n, that q_{n+1} is orthogonal to q_n. ■

The following two theorems tell us more about the Legendre polynomials. First we show that the recurrence relation in (2.18) can be simplified.

Theorem 2.1.2 The Legendre polynomials, scaled so that their leading coefficients are unity, satisfy the recurrence relation

$$q_{n+1}(x) = xq_n(x) - \beta_n q_{n-1}(x), \tag{2.21}$$

for $n \geq 1$, with $q_0(x) = 1$ and $q_1(x) = x$, where β_n is defined by (2.20). Further, q_n is an even function when n is even, and is an odd function when n is odd.

Proof. We begin by noting that $q_0(x) = 1$ is an even function, and $q_1(x) = x$ is odd. Let us assume that for some $k \geq 0$, the polynomials $q_0, q_1, \ldots, q_{2k+1}$ are alternately even and odd, and that $\alpha_n = 0$ in (2.19) for all $n \leq 2k$. It then follows from (2.19) that $\alpha_{2k+1} = 0$, and then from the recurrence relation (2.18) that q_{2k+2} is an even function. Another inspection of (2.19) shows that $\alpha_{2k+2} = 0$, and then (2.18) shows that q_{2k+3} is an odd function. The proof is completed by induction. Later in this section we will determine the value of β_n. ■

Theorem 2.1.3 The Legendre polynomial of degree n has n distinct zeros in the interior of the interval $[-1, 1]$.

Proof. Since q_n is orthogonal to 1 for $n > 0$,

$$\int_{-1}^{1} q_n(x)dx = 0, \quad n > 0,$$

and thus q_n must have at least one zero in the interior of $[-1, 1]$. If $x = x_1$ were a multiple zero of q_n, for $n \geq 2$, then $q_n(x)/(x - x_1)^2$ would be a polynomial in P_{n-2} and so be orthogonal to q_n, giving

$$0 = \int_{-1}^{1} \frac{q_n(x)}{(x - x_1)^2}\, q_n(x)dx = \int_{-1}^{1} \left(\frac{q_n(x)}{x - x_1}\right)^2 dx,$$

which is impossible. Thus the zeros of q_n are all distinct. Now suppose that q_n has exactly $k \geq 1$ zeros in the interior of $[-1, 1]$, and that

$$q_n(x) = (x - x_1) \cdots (x - x_k)\, r(x) = \pi_k(x) r(x),$$

say, where $r(x)$ does not change sign in $(-1, 1)$. Then if $k < n$, it follows from the orthogonality property that

$$0 = \int_{-1}^{1} \pi_k(x) q_n(x)dx = \int_{-1}^{1} [\pi_k(x)]^2 r(x)dx,$$

which is impossible, since $[\pi_k(x)]^2 r(x)$ does not change sign. Thus we must have $k = n$, and consequently $r(x) = 1$, which completes the proof. ■

We will now obtain an explicit form for the recurrence relation (2.21) for the Legendre polynomials. Consider derivatives of the function

$$(x^2-1)^n = (x-1)^n(x+1)^n.$$

It is easily verified, using the Leibniz rule for differentiation (1.83), that

$$\frac{d^j}{dx^j}(x^2-1)^n = 0, \quad 0 \le j \le n-1, \quad \text{for} \quad x = \pm 1, \tag{2.22}$$

and

$$\frac{d^n}{dx^n}(x^2-1)^n = 2^n n! \quad \text{for} \quad x = 1. \tag{2.23}$$

If we define

$$Q_n(x) = \frac{1}{2^n n!}\,\frac{d^n}{dx^n}(x^2-1)^n, \tag{2.24}$$

it is clear that Q_n is a polynomial of degree n, and that $Q_n(1) = 1$. The relation (2.24) is called a Rodrigues formula, after O. Rodrigues. We now state and prove the following lemma.

Lemma 2.1.1 If u and v are both n times differentiable on $[-1,1]$, and v and its first $n-1$ derivatives are zero at both endpoints $x = \pm 1$, then

$$\int_{-1}^{1} u(x)v^{(n)}(x)dx = (-1)^n \int_{-1}^{1} u^{(n)}(x)v(x)dx, \quad n \ge 1. \tag{2.25}$$

Proof. Using integration by parts, we have

$$\int_{-1}^{1} u(x)v^{(n)}(x)dx = -\int_{-1}^{1} u'(x)v^{(n-1)}(x)dx,$$

and the proof is completed by using induction on n. ■

In particular, if g is any function that is n times differentiable, we obtain

$$\int_{-1}^{1} g(x)\frac{d^n}{dx^n}(x^2-1)^n dx = \int_{-1}^{1} g^{(n)}(x)(1-x^2)^n dx, \tag{2.26}$$

and the latter integral is zero if $g \in P_{n-1}$. We deduce that the polynomial $Q_n \in P_n$ is orthogonal to all polynomials in P_{n-1}, and thus must be a multiple of the Legendre polynomial q_n. It can be shown that

$$|Q_n(x)| \le 1, \quad \text{for} \quad |x| \le 1, \tag{2.27}$$

the maximum modulus of 1 being attained at the endpoints $x = \pm 1$. (This inequality is derived in Rivlin [48], via a cleverly arranged sequence of results involving Q_n and its first derivative.) To derive an explicit form of the recurrence relation (2.21), we write

$$(x^2-1)^n = x^{2n} - nx^{2n-2} + \cdots,$$

and so obtain from (2.24) that

$$Q_n(x) = \frac{1}{2^n n!}\left(\frac{(2n)!}{n!}x^n - \frac{n(2n-2)!}{(n-2)!}x^{n-2} + \cdots\right), \tag{2.28}$$

for $n \geq 2$. Since the Legendre polynomial q_n has leading term x^n, with coefficient unity, it follows from (2.28) that

$$Q_n(x) = \mu_n q_n(x), \quad \text{where} \quad \mu_n = \frac{1}{2^n}\binom{2n}{n}, \tag{2.29}$$

and

$$q_n(x) = x^n - \frac{n(n-1)}{2(2n-1)}x^{n-2} + \cdots. \tag{2.30}$$

If we now use (2.30) in the recurrence relation (2.21), we obtain

$$x^{n+1} - \frac{(n+1)n}{2(2n+1)}x^{n-1} + \cdots = x^{n+1} - \frac{n(n-1)}{2(2n-1)}x^{n-1} - \beta_n x^{n-1} + \cdots,$$

and on equating coefficients of x^{n-1}, we find that

$$\beta_n = \frac{(n+1)n}{2(2n+1)} - \frac{n(n-1)}{2(2n-1)} = \frac{n^2}{4n^2-1}. \tag{2.31}$$

It was convenient to work with q_n, which is scaled so that its coefficient of x^n is unity, in the early part of our discussion on the Legendre polynomials. This enabled us to simplify the recurrence relation given in Theorems 2.1.1 and 2.1.2. In order to refine the recurrence relation, we then introduced Q_n, the multiple of q_n scaled so that $Q_n(1) = 1$. In the mathematical literature it is Q_n that is called *the* Legendre polynomial. If we use (2.29) to replace each q_n in the recurrence relation (2.21) by the appropriate multiple of Q_n, and use (2.31), we obtain the following remarkably simple recurrence relation, which deserves to be expressed as a theorem.

Theorem 2.1.4 The Legendre polynomials Q_n satisfy the recurrence relation

$$(n+1)Q_{n+1}(x) = (2n+1)xQ_n(x) - nQ_{n-1}(x), \tag{2.32}$$

for $n \geq 1$, where $Q_0(x) = 1$ and $Q_1(x) = x$. ■

We find from the recurrence relation (2.32) and its initial conditions that the first few Legendre polynomials are

$$1, \quad x, \quad \frac{1}{2}(3x^2-1), \quad \frac{1}{2}(5x^3-3x), \quad \frac{1}{8}(35x^4-30x^2+3). \tag{2.33}$$

The Legendre polynomial $Q_n(x)$ (see Problem 2.1.7) satisfies the second-order differential equation

$$(1-x^2)Q_n''(x) - 2xQ_n'(x) + n(n+1)Q_n(x) = 0. \tag{2.34}$$

Let us return to a problem that we posed near the beginning of this section: Given a function f defined on $[-1,1]$, let us seek the polynomial $p_n \in P_n$ that minimizes $\|f - p_n\|$, where $\|\cdot\|$ is the square norm. Our solution to this problem was given in terms of the polynomials q_r, and if we recast this in terms of the Legendre polynomials proper, we obtain the solution

$$p_n(x) = \sum_{r=0}^{n} a_r Q_r(x), \tag{2.35}$$

where

$$a_r = \int_{-1}^{1} f(x)Q_r(x)dx \Big/ \int_{-1}^{1} [Q_r(x)]^2 dx, \quad 0 \le r \le n. \tag{2.36}$$

We call the polynomial p_n the best square norm approximation or, more commonly, the *least squares* approximation for f on $[-1,1]$. We remark, in passing, that the partial sum of a Fourier series has this same property of being a least squares approximation. If we let $n \to \infty$ in (2.35), the resulting infinite series, if it exists, is called the *Legendre series* for f. From (2.36) and Problem 2.1.6, the Legendre coefficients may be expressed as

$$a_r = \frac{1}{2}(2r+1)\int_{-1}^{1} f(x)Q_r(x)dx, \quad r \ge 0, \tag{2.37}$$

and we have the following result.

Theorem 2.1.5 The partial sum of the Legendre series for f is even or odd if f is even or odd, respectively.

Proof. As we saw in Theorem 2.1.2, the polynomial q_n is an even function when n is even, and is an odd function when n is odd, and (2.29) shows that this holds also for its multiple, the Legendre polynomial Q_n. It follows from (2.37) that the Legendre coefficient a_r is zero if r is odd and f is even, and is also zero if r is even and f is odd. Thus the Legendre series for f contains only even- or odd-order Legendre polynomials when f is even or odd, respectively. This completes the proof. ■

If f is sufficiently differentiable, we can derive another expression for the Legendre coefficients by expressing $Q_s(x)$ in (2.37) in its Rodrigues form, given in (2.24), and then use (2.26) to give

$$a_r = \frac{2r+1}{2^{r+1}r!}\int_{-1}^{1} f^{(r)}(x)(1-x^2)^r dx. \tag{2.38}$$

We now give an estimate of the size of the Legendre coefficient a_r for f in terms of $f^{(r)}$.

Theorem 2.1.6 If $f^{(r)}$ is continuous on $[-1, 1]$, the Legendre coefficient a_r is given by

$$a_r = \frac{2^r\, r!}{(2r)!}\, f^{(r)}(\xi_r), \tag{2.39}$$

where $\xi_r \in (-1, 1)$.

Proof. Since $f^{(r)}$ is continuous on $[-1, 1]$, it follows from the mean value theorem for integrals, Theorem 3.1.2, that

$$a_r = \frac{2r+1}{2^{r+1} r!} f^{(r)}(\xi_r) I_r,$$

where

$$I_r = \int_{-1}^{1} (1 - x^2)^r dx.$$

Then, using integration by parts, we find that

$$I_r = 2r \int_{-1}^{1} x^2 (1 - x^2)^{r-1} dx = 2r(I_{r-1} - I_r).$$

Hence

$$I_r = \frac{2r}{2r+1} I_{r-1} = \frac{2r}{2r+1} \frac{2r-2}{2r-1} \cdots \frac{2}{3} I_0,$$

where $I_0 = 2$, and (2.39) follows easily. ■

Example 2.1.2 Let us use (2.38) to compute the Legendre coefficients for the function e^x. First we derive from (2.38)

$$a_0 = \frac{1}{2}(e - e^{-1}) \approx 1.175201, \qquad a_1 = 3e^{-1} \approx 1.103638.$$

If we write

$$J_r = \frac{1}{2^r\, r!} \int_{-1}^{1} e^x (1 - x^2)^r dx, \tag{2.40}$$

then (2.38) with $f(x) = e^x$ yields $a_r = \frac{1}{2}(2r+1)J_r$. On using integration by parts twice on the above integral for J_r, we obtain

$$J_r = -(2r - 1)J_{r-1} + J_{r-2},$$

and hence obtain the recurrence relation

$$\frac{a_r}{2r+1} = -a_{r-1} + \frac{a_{r-2}}{2r-3}, \tag{2.41}$$

with a_0 and a_1 as given above. The next few Legendre coefficients for e^x, rounded to six decimal places, are as follows:

n	2	3	4	5	6	7
a_n	0.357814	0.070456	0.009965	0.001100	0.000099	0.000008

All the Legendre coefficients for e^x are positive. (See Problem 2.1.10.) Before leaving this example, we remark that the recurrence relation (2.41) is numerically unstable, since the error in the computed value of a_r is approximately $2r + 1$ times the error in a_{r-1}. Thus, if we need to compute a_r for a large value of r, we should estimate it directly from the expression (2.40) for J_r, using a numerical integration method. ■

It is convenient to write

$$(f, g) = \int_{-1}^{1} f(x)g(x)dx, \tag{2.42}$$

and we call (f, g) an *inner product* of f and g. It follows from (2.35), the orthogonality of the Legendre polynomials Q_s, and (2.36) that

$$(p_n, Q_s) = a_s(Q_s, Q_s) = (f, Q_s). \tag{2.43}$$

Since

$$(f - p_n, Q_s) = (f, Q_s) - (p_n, Q_s),$$

we see from (2.43) that

$$(f - p_n, Q_s) = 0, \quad 0 \leq s \leq n. \tag{2.44}$$

For further material on inner products, see Davis [10], Deutsch [14]. The following theorem describes another property concerning the difference between f and p_n.

Theorem 2.1.7 The error term $f - p_n$ changes sign on at least $n+1$ points in the interior of $[-1, 1]$, where p_n is the partial sum of the Legendre series for f.

Proof. First we have from (2.44) that $(f - p_n, Q_0) = 0$. Since $Q_0(x) = 1$, it follows that there is at least one point in $(-1, 1)$ where $f - p_n$ changes sign. Suppose that $f(x) - p_n(x)$ changes sign at k points,

$$-1 < x_1 < x_2 < \cdots < x_k < 1,$$

and at no other points in $(-1, 1)$, with $1 \leq k < n+1$. Then $f(x) - p_n(x)$ and the function $\pi_k(x) = (x - x_1) \cdots (x - x_k)$ change sign at the x_j, and at no other points in $(-1, 1)$. Thus we must have $(f - p_n, \pi_k) \neq 0$. On the other hand, since π_k may be written as a sum of multiples of $Q_0, Q_1, \ldots, Q_k$, it follows from (2.44) that $(f - p_n, \pi_k) = 0$, which gives a contradiction. We deduce that $k \geq n + 1$, which completes the proof. ■

Since, as we have just established, there are at least $n + 1$ points in $(-1, 1)$ where $p_n(x)$ and $f(x)$ are equal, the best approximant p_n must be an interpolating polynomial for f. This observation leads us to the following interesting expression for $\|f - p_n\|$, of a similar form to the error term for the Taylor polynomial or the error term for the interpolating polynomial, which are given in (1.35) and (1.25), respectively.

Theorem 2.1.8 Let p_n denote the least squares approximation for f on $[-1,1]$. Then if $f^{(n+1)}$ is continuous on $[-1,1]$ there exists a number ζ in $(-1,1)$ such that

$$\|f - p_n\| = \frac{1}{\mu_{n+1}}\sqrt{\frac{2}{2n+3}}\,\frac{|f^{(n+1)}(\zeta)|}{(n+1)!} \sim \frac{\sqrt{\pi}}{2^{n+1}}\,\frac{|f^{(n+1)}(\zeta)|}{(n+1)!}, \tag{2.45}$$

where μ_{n+1} is defined in (2.29).

Proof. We begin with a comment on the notation used in (2.45). We write $u_n \sim v_n$ to mean that

$$\lim_{n\to\infty} \frac{u_n}{v_n} = 1,$$

and we say that u_n and v_n are asymptotically equal, as $n \to \infty$.

Since the best approximant p_n is an interpolating polynomial for f, we have from (1.25) that

$$f(x) - p_n(x) = (x - x_0)(x - x_1)\cdots(x - x_n)\frac{f^{n+1}(\xi_x)}{(n+1)!},$$

where $\xi_x \in (-1,1)$, and the x_j are distinct points in $(-1,1)$. Taking the square norm, and applying the mean value theorem for integrals, Theorem 3.1.2, we find that

$$\|f - p_n\| = \frac{|f^{(n+1)}(\xi)|}{(n+1)!}\,\|(x - x_0)\cdots(x - x_n)\|, \tag{2.46}$$

for some $\xi \in (-1,1)$. If p_n^* denotes the interpolating polynomial for f on the zeros of the Legendre polynomial Q_{n+1}, we similarly have

$$\|f - p_n^*\| = \frac{|f^{(n+1)}(\eta)|}{(n+1)!}\,\|(x - x_0^*)\cdots(x - x_n^*)\|, \tag{2.47}$$

where $\eta \in (-1,1)$, and the x_j^* are the zeros of Q_{n+1}. We then see from (2.47) and Problem 2.1.9 that

$$\|f - p_n^*\| = \frac{1}{\mu_{n+1}}\sqrt{\frac{2}{2n+3}}\,\frac{|f^{(n+1)}(\eta)|}{(n+1)!},$$

where μ_{n+1} is defined in (2.29). Since p_n is the least squares approximation for f, we have

$$\|f - p_n\| \le \|f - p_n^*\| = \frac{1}{\mu_{n+1}}\sqrt{\frac{2}{2n+3}}\,\frac{|f^{(n+1)}(\eta)|}{(n+1)!}, \tag{2.48}$$

where $\eta \in (-1,1)$. Also, we may deduce from (2.46) and Problem 2.1.9 that

$$\|f - p_n\| \ge \frac{1}{\mu_{n+1}}\sqrt{\frac{2}{2n+3}}\,\frac{|f^{(n+1)}(\xi)|}{(n+1)!}, \tag{2.49}$$

where $\xi \in (-1,1)$. We can now combine (2.48) and (2.49), and use the continuity of $f^{(n+1)}$ to give

$$\|f - p_n\| = \frac{1}{\mu_{n+1}} \sqrt{\frac{2}{2n+3}} \frac{|f^{(n+1)}(\zeta)|}{(n+1)!}, \quad \zeta \in (-1,1),$$

and we complete the proof by applying Stirling's formula (see Problem 2.1.12) to give

$$\frac{1}{\mu_{n+1}} \sqrt{\frac{2}{2n+3}} \sim \frac{\sqrt{\pi}}{2^n}. \qquad \blacksquare$$

Problem 2.1.1 Let

$$\|x\| = \left(x_1^2 + \cdots + x_n^2\right)^{1/2}.$$

Verify the inequality

$$(x_1 y_1 + \cdots + x_n y_n)^2 \le (x_1^2 + \cdots + x_n^2)(y_1^2 + \cdots + y_n^2)$$

by showing that it is equivalent to

$$\sum_{i \neq j} (x_i y_j - x_j y_i)^2 \ge 0,$$

where the latter sum of $\frac{1}{2}n(n-1)$ terms is taken over all distinct pairs of numbers i and j chosen from the set $\{1, \ldots, n\}$. Deduce the inequality

$$|x_1 y_1 + \cdots + x_n y_n| \le \|x\| \cdot \|y\|,$$

and hence justify the triangle inequality for this norm.

Problem 2.1.2 Define

$$f(x) = \begin{cases} 0, & -1 \le x < 0, \\ x, & 0 \le x \le 1, \end{cases}$$

and define $g(x) = f(-x)$, $-1 \le x \le 1$. Show that

$$\|f\| = \|g\| = \frac{1}{(p+1)^{1/p}} \quad \text{and} \quad \|f + g\| = \frac{2^{1/p}}{(p+1)^{1/p}},$$

where $\|\cdot\|$ denotes the p-norm, defined by (2.7). Deduce that $p \ge 1$ is a necessary condition for the triangle inequality to hold.

Problem 2.1.3 Scale the Legendre polynomials 1, x, $\frac{1}{2}(3x^2 - 1)$, and $\frac{1}{2}(5x^3 - 3x)$ so that they form an orthonormal basis for P_3.

Problem 2.1.4 Verify (2.22) and (2.23).

Problem 2.1.5 By writing down the binomial expansion of $(x^2-1)^n$ and differentiating it n times, deduce from (2.24) that

$$Q_n(x) = \frac{1}{2^n} \sum_{j=0}^{[n/2]} (-1)^j \left(\begin{array}{c} n \\ j \end{array} \right) \left(\begin{array}{c} 2n-2j \\ n \end{array} \right) x^{n-2j},$$

where $[n/2]$ denotes the integer part of $n/2$.

Problem 2.1.6 We have, from (2.20) and (2.31),

$$\beta_n = \int_{-1}^{1} [q_n(x)]^2\, dx \Big/ \int_{-1}^{1} [q_{n-1}(x)]^2\, dx = \frac{n^2}{4n^2-1},$$

for $n \geq 1$. Use this and (2.29) to deduce that

$$\int_{-1}^{1} [Q_n(x)]^2\, dx \Big/ \int_{-1}^{1} [Q_{n-1}(x)]^2\, dx = \frac{2n-1}{2n+1},$$

for $n \geq 1$, and hence show that

$$\int_{-1}^{1} [Q_n(x)]^2\, dx = \frac{2}{2n+1}$$

for $n \geq 0$.

Problem 2.1.7 If $y(x)$ and $p(x)$ are functions that are twice differentiable, use integration by parts twice to show that

$$\begin{aligned} \int_{-1}^{1} \frac{d}{dx}\{(1-x^2)y'(x)\}\, p(x)dx &= -\int_{-1}^{1} (1-x^2)\, y'(x)p'(x)dx \\ &= \int_{-1}^{1} \frac{d}{dx}\{(1-x^2)p'(x)\}\, y(x)dx. \end{aligned}$$

Now write $y(x) = Q_n(x)$ and let $p \in P_{n-1}$, and deduce that

$$\frac{d}{dx}\{(1-x^2)Q_n'(x)\} = (1-x^2)Q_n''(x) - 2xQ_n'(x)$$

is orthogonal to all polynomials in P_{n-1}, and thus must be a multiple of the Legendre polynomial Q_n. Using (2.28), equate coefficients of x^n in

$$(1-x^2)Q_n''(x) - 2xQ_n'(x) = \mu Q_n(x),$$

to give

$$-n(n-1) - 2n = \mu,$$

and thus show that Q_n satisfies the second-order differential equation given in (2.34).

Problem 2.1.8 Show that the polynomial $p \in P_n$ that minimizes

$$\int_{-1}^{1} [x^{n+1} - p(x)]^2 dx$$

is $p(x) = x^{n+1} - q_{n+1}(x)$, where q_{n+1} is the Legendre polynomial Q_{n+1} scaled so that it has leading coefficient unity. Hint: Write

$$x^{n+1} - p(x) = q_{n+1}(x) + \sum_{j=0}^{n} \gamma_j q_j(x),$$

and use the orthogonality properties of the q_j.

Problem 2.1.9 Deduce from the result in the previous problem that for the square norm on $[-1, 1]$, the minimum value of $\|(x - x_0) \cdots (x - x_n)\|$ is attained when the x_j are the zeros of the Legendre polynomial Q_{n+1}, and use the result in Problem 2.1.6 to show that

$$\min_{x_j} \|(x - x_0) \cdots (x - x_n)\| = \frac{1}{\mu_{n+1}} \sqrt{\frac{2}{2n+3}}\,,$$

where μ_{n+1} is defined in (2.29).

Problem 2.1.10 Deduce from (2.38) that if f and all its derivatives are nonnegative on $[-1, 1]$, then all coefficients of the Legendre series for f are nonnegative.

Problem 2.1.11 Verify that

$$\frac{d}{dx}(x^2 - 1)^n = 2nx(x^2 - 1)^{n-1},$$

differentiate n times, and use the Leibniz rule (1.83) to show that

$$\frac{d^{n+1}}{dx^{n+1}}(x^2 - 1)^n = 2n \left(x \frac{d^n}{dx^n}(x^2 - 1)^{n-1} + n \frac{d^{n-1}}{dx^{n-1}}(x^2 - 1)^{n-1} \right),$$

for $n \geq 1$. Finally, divide by $2^n n!$ and use the Rodrigues formula (2.24) to obtain the relation

$$Q_n'(x) = xQ_{n-1}'(x) + nQ_{n-1}(x), \quad n \geq 1.$$

Problem 2.1.12 Use Stirling's formula,

$$n! \sim \sqrt{2\pi n} \left(\frac{n}{e}\right)^n,$$

to show that (see (2.29))

$$\mu_n = \frac{1}{2^n} \binom{2n}{n} \sim \frac{2^n}{\sqrt{\pi n}}.$$

2.2 The Chebyshev Polynomials

Many of the ideas and formulas presented in the last section concerning the Legendre polynomials can be generalized by introducing a *weight function.* This generalization leads to an infinite number of systems of orthogonal polynomials. We will pay particular attention to one such system, the Chebyshev polynomials, named after P. L. Chebyshev (1821–1894).

Given any integrable function ω that is nonnegative and not identically zero on $[-1,1]$, we can construct a sequence of polynomials (q_n^{ω}), where q_n^{ω} is of degree n, has leading coefficient unity, and satisfies

$$\int_{-1}^{1} \omega(x) x^r q_n^{\omega}(x) dx = 0, \qquad 0 \le r < n. \tag{2.50}$$

The polynomials q_n^{ω} are said to be orthogonal on $[-1,1]$ with respect to the weight function ω, and the scaled Legendre polynomials q_n are recovered by putting $\omega(x)=1$. The generalized orthogonal polynomials q_n^{ω}, like the Legendre polynomials, satisfy a recurrence relation of the form (2.18), where the coefficients α_n and β_n are given by

$$\alpha_n = \int_{-1}^{1} \omega(x) x [q_n^{\omega}(x)]^2 dx \Big/ \int_{-1}^{1} \omega(x) [q_n^{\omega}(x)]^2 dx \tag{2.51}$$

and

$$\beta_n = \int_{-1}^{1} \omega(x) [q_n^{\omega}(x)]^2 dx \Big/ \int_{-1}^{1} \omega(x) [q_{n-1}^{\omega}(x)]^2 dx. \tag{2.52}$$

Further, if the weight function ω is even, then $\alpha_n = 0$, the even-order orthogonal polynomials are even functions, and the odd-order polynomials are odd, as given by Theorem 2.1.2 for the Legendre polynomials. The above statements about the generalized orthogonal polynomials are easily verified by inserting the weight function ω appropriately and repeating the arguments used above in the special case where $\omega(x)=1$. See, for example, Davis and Rabinowitz [11].

Let us now define

$$\|f\| = \left(\int_{-1}^{1} \omega(x) [f(x)]^2 dx \right)^{1/2}. \tag{2.53}$$

It can be shown that $\|\cdot\|$ in (2.53) satisfies the three properties (2.2), (2.3), and (2.4), and so indeed defines a norm, which we call a weighted square norm. When we choose $\omega(x)=1$, we recover the square norm (2.1). We then find that

$$\|f-p\| = \left(\int_{-1}^{1} \omega(x) [f(x)-p(x)]^2 dx \right)^{1/2}$$

is minimized over all $p \in P_n$ by choosing $p = p_n$, where

$$p_n(x) = \sum_{r=0}^{n} a_r q_r^\omega(x), \tag{2.54}$$

and each coefficient a_r is given by

$$a_r = \int_{-1}^{1} \omega(x) f(x) q_r^\omega(x) dx \bigg/ \int_{-1}^{1} \omega(x) [q_r^\omega(x)]^2 dx, \quad 0 \leq r \leq n. \tag{2.55}$$

If we let $n \to \infty$, we obtain a generalized orthogonal expansion for f whose coefficients are given by (2.55). In particular, if $\omega(x) = 1$, we obtain the Legendre series. We can easily adapt Theorem 2.1.5 to show that if the weight function ω is even, then the partial sum of the generalized orthogonal series for f is even or odd if f is even or odd, respectively.

The choice of weight function

$$\omega(x) = (1-x)^\alpha (1+x)^\beta, \quad \alpha, \beta > -1, \tag{2.56}$$

leads to a two-parameter system of orthogonal polynomials that are called the Jacobi polynomials. These include the following, as special cases:

$\alpha = \beta = 0$	Legendre polynomials,
$\alpha = \beta = -\frac{1}{2}$	Chebyshev polynomials,
$\alpha = \beta = \frac{1}{2}$	Chebyshev polynomials of the second kind,
$\alpha = \beta$	ultraspherical polynomials.

Note that the first three systems of orthogonal polynomials listed above are all special cases of the fourth system, the *ultraspherical* polynomials, whose weight function is the even function $(1-x^2)^\alpha$. Thus the ultraspherical polynomials satisfy a recurrence relation of the form given in Theorem 2.1.2 for the Legendre polynomials. Further, the ultraspherical polynomial of degree n is even or odd, when n is even or odd, respectively. If we choose ω as the Jacobi weight function (2.56), then (2.54) and (2.55) define a partial Jacobi series, and it is clear that when $\alpha = \beta$, the resulting partial ultraspherical series is even or odd if f is even or odd, respectively.

To investigate the Jacobi polynomials, we begin with the function

$$(1-x)^{-\alpha}(1+x)^{-\beta} \frac{d^n}{dx^n}(1-x)^{n+\alpha}(1+x)^{n+\beta}. \tag{2.57}$$

First, using the Leibniz rule (1.83) for differentiating a product, we can show that this function is a polynomial of degree n. Then, following the method we used for the Legendre polynomials, beginning with Lemma

2.1.1, we can show that this polynomial is orthogonal on $[-1, 1]$, with respect to the Jacobi weight function given in (2.56), to all polynomials in P_{n-1}. We then define the Jacobi polynomial of degree n as

$$Q_n^{(\alpha,\beta)}(x) = \frac{(-1)^n}{2^n n!}(1-x)^{-\alpha}(1+x)^{-\beta}\frac{d^n}{dx^n}(1-x)^{n+\alpha}(1+x)^{n+\beta}. \quad (2.58)$$

It is easy to adapt the proof of Theorem 2.1.3 to show that the Jacobi polynomial of degree n has n distinct zeros in the interior of $[-1, 1]$.

The following theorem generalizes Theorem 2.1.4.

Theorem 2.2.1 The Jacobi polynomials satisfy the recurrence relation

$$Q_{n+1}^{(\alpha,\beta)}(x) = (a_n x + b_n)Q_n^{(\alpha,\beta)}(x) - c_n Q_{n-1}^{(\alpha,\beta)}(x), \quad (2.59)$$

where

$$a_n = \frac{(2n+\alpha+\beta+1)(2n+\alpha+\beta+2)}{2(n+1)(n+\alpha+\beta+1)},$$
$$b_n = \frac{(\alpha^2-\beta^2)(2n+\alpha+\beta+1)}{2(n+1)(n+\alpha+\beta+1)(2n+\alpha+\beta)},$$
$$c_n = \frac{(n+\alpha)(n+\beta)(2n+\alpha+\beta+2)}{(n+1)(n+\alpha+\beta+1)(2n+\alpha+\beta)}.$$

Proof. This can be justified by following the same method as we used to prove Theorem 2.1.4. See Davis [10]. ■

We can use the Leibniz rule (1.83) again to show that

$$Q_n^{(\alpha,\beta)}(1) = \frac{1}{n!}(n+\alpha)(n-1+\alpha)\cdots(1+\alpha) = \binom{n+\alpha}{n}, \quad (2.60)$$

which is independent of β. Likewise (see Problem 2.2.2), we can show that $Q_n^{(\alpha,\beta)}(-1)$ is independent of α. We can express (2.60) in the form

$$Q_n^{(\alpha,\beta)}(1) = \frac{\Gamma(n+\alpha+1)}{\Gamma(n+1)\Gamma(\alpha+1)}, \quad (2.61)$$

where Γ is the gamma function, defined by

$$\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}dt. \quad (2.62)$$

It is not hard (see Problem 2.2.1) to deduce from (2.62) that the gamma function satisfies the difference equation

$$\Gamma(x+1) = x\Gamma(x), \quad (2.63)$$

with $\Gamma(1) = 1$, and hence justify (2.61).

From (2.55) the Jacobi series has coefficients that satisfy

$$a_r = \int_{-1}^{1} \omega(x) f(x) Q_r^{(\alpha,\beta)}(x) dx \Big/ \int_{-1}^{1} \omega(x) \left[Q_r^{(\alpha,\beta)}(x) \right]^2 dx, \tag{2.64}$$

where $\omega(x) = (1-x)^\alpha (1+x)^\beta$. Then, if f is sufficiently differentiable, it follows from (2.58) and Lemma 2.1.1 that the numerator on the right of (2.64) may be written as

$$\int_{-1}^{1} \omega(x) f(x) Q_r^{(\alpha,\beta)}(x) dx = \frac{1}{2^r\, r!} \int_{-1}^{1} \omega(x) f^{(r)}(x) (1-x^2)^r dx, \tag{2.65}$$

where $\omega(x) = (1-x)^\alpha (1+x)^\beta$. The denominator on the right of (2.64) can be expressed (see Davis [10]) in the form

$$\int_{-1}^{1} \omega(x) [Q_r^{(\alpha,\beta)}(x)]^2 dx = \frac{2^{\alpha+\beta+1}}{(2r+\alpha+\beta+1)} \frac{\Gamma(r+\alpha+1)\Gamma(r+\beta+1)}{\Gamma(r+1)\Gamma(r+\alpha+\beta+1)},$$

and we note that this is consistent with our findings in Problem 2.1.6 for the special case of the Legendre polynomials.

One might argue that the simplest of all the Jacobi polynomials are the Legendre polynomials, since these have the simplest weight function. However, there are good reasons for saying that the simplest Jacobi polynomials are the Chebyshev polynomials, which have weight function $(1-x^2)^{-1/2}$. Note that the latter weight function is singular at the endpoints $x = \pm 1$. The Chebyshev polynomials are usually denoted by $T_n(x)$, and are uniquely defined by

$$\int_{-1}^{1} (1-x^2)^{-1/2} T_r(x) T_s(x) dx = 0, \quad r \neq s, \tag{2.66}$$

where

$$T_r \in P_r \quad \text{and} \quad T_r(1) = 1, \quad r \geq 0. \tag{2.67}$$

It follows from the above definition of the Chebyshev polynomials that T_n is a multiple of the Jacobi (also ultraspherical) polynomial $Q_n^{(-1/2,-1/2)}$. With $\alpha = \beta = -\frac{1}{2}$ in (2.60), we can show that

$$Q_n^{(-1/2,-1/2)}(1) = \frac{1}{2^{2n}} \binom{2n}{n} \sim \frac{1}{\sqrt{\pi n}}, \tag{2.68}$$

and, since $T_n(1) = 1$, it follows that

$$Q_n^{(-1/2,-1/2)}(x) = \frac{1}{2^{2n}} \binom{2n}{n} T_n(x). \tag{2.69}$$

The following theorem shows that the Chebyshev polynomials can be expressed in a very simple form in terms of the cosine function.

Theorem 2.2.2 Let us define

$$t_n(x) = \cos n\theta, \quad \text{where } x = \cos\theta, \quad -1 \le x \le 1, \quad n \ge 0. \tag{2.70}$$

Then $t_n = T_n$, the Chebyshev polynomial of degree n.

Proof. On making the substitution $x = \cos\theta$, the interval $-1 \le x \le 1$ corresponds to $0 \le \theta \le \pi$, and we obtain

$$\int_{-1}^{1} (1-x^2)^{-1/2} t_r(x) t_s(x) dx = \int_0^{\pi} \cos r\theta \cos s\theta \, d\theta, \quad r \ne s.$$

Since

$$\cos r\theta \cos s\theta = \frac{1}{2}(\cos(r+s)\theta + \cos(r-s)\theta),$$

we readily see that

$$\int_{-1}^{1} (1-x^2)^{-1/2} t_r(x) t_s(x) dx = 0, \quad r \ne s.$$

From the substitution $x = \cos\theta$, we see that $x = 1$ corresponds to $\theta = 0$, and thus $t_r(1) = 1$. Now

$$\cos(n+1)\theta + \cos(n-1)\theta = 2\cos n\theta \cos\theta,$$

which, using (2.70), yields the recurrence relation

$$t_{n+1}(x) = 2x\, t_n(x) - t_{n-1}(x), \quad n \ge 1,$$

and we see from (2.70) that $t_0(x) = 1$ and $t_1(x) = x$. Thus, by induction, each t_r belongs to P_r. It follows that $t_r = T_r$ for all $r \ge 0$, and this completes the proof. ■

We have just shown that

$$T_n(x) = \cos n\theta, \quad \text{where } x = \cos\theta, \quad -1 \le x \le 1, \quad n \ge 0, \tag{2.71}$$

and that the Chebyshev polynomials satisfy the recurrence relation

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x), \quad n \ge 1, \tag{2.72}$$

with $T_0(x) = 1$ and $T_1(x) = x$. We find that the first few Chebyshev polynomials are

$$1, \quad x, \quad 2x^2 - 1, \quad 4x^3 - 3x, \quad 8x^4 - 8x^2 + 1. \tag{2.73}$$

Given how simply the Chebyshev polynomial is expressed in (2.71), we can very easily find its zeros and turning values. We write

$$T_n(x) = 0 \quad \Rightarrow \quad \cos n\theta = 0 \quad \Rightarrow \quad n\theta = (2j-1)\frac{\pi}{2}, \tag{2.74}$$

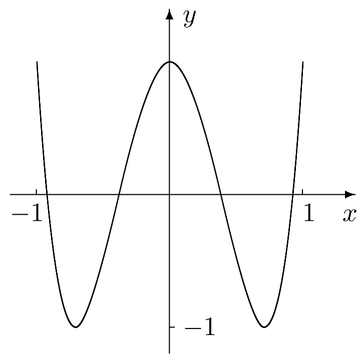

FIGURE 2.1. Graph of the Chebyshev polynomial $T_4(x)$.

where j is an integer. Thus

$$T_n(x) = 0 \quad \Rightarrow \quad x = \cos\theta_j, \qquad \text{where} \quad \theta_j = \frac{(2j-1)\pi}{2n}, \tag{2.75}$$

for some integer j. We see from the graph of $x = \cos\theta$ that as θ takes all values between 0 and π, the function $x = \cos\theta$ is monotonic decreasing, and takes all values between 1 and -1. The choice of $j = 1, 2, \ldots, n$ in (2.75) gives n distinct zeros of T_n in $(-1, 1)$, and since $T_n \in P_n$, *all* the zeros of T_n are given by the n values

$$x_j = \cos\frac{(2j-1)\pi}{2n}, \quad 1 \le j \le n. \tag{2.76}$$

Since for $x \in [-1, 1]$ we can express $T_n(x)$ in the form $\cos n\theta$, where $x = \cos\theta$, it is clear that the maximum modulus of T_n on $[-1, 1]$ is 1. This is attained for values of θ such that $|\cos n\theta| = 1$, and

$$\cos n\theta = \pm 1 \quad \Rightarrow \quad n\theta = j\pi \quad \Rightarrow \quad x = \cos(j\pi/n) = \tau_j,$$

say, where j is an integer. Thus the Chebyshev polynomial T_n attains its maximum modulus of 1 at the $n+1$ points $\tau_j = \cos(j\pi/n)$, for $0 \le j \le n$, and T_n alternates in sign over this set of points. For we have

$$T_n(\tau_j) = \cos j\pi = (-1)^j, \quad 0 \le j \le n.$$

These $n + 1$ points of maximum modulus are called the *extreme points* of T_n. The Chebyshev polynomial T_n is the only polynomial in P_n whose maximum modulus is attained on $n + 1$ points of $[-1, 1]$. Note that although we have expressed $T_n(x)$ in terms of the cosine function for the interval $-1 \le x \le 1$ only, the Chebyshev polynomial is defined by its recurrence relation (2.72) for all real x. Outside the interval $[-1, 1]$, we can (see Problem 2.2.6) express $T_n(x)$ in terms of the hyperbolic cosine. For further information on the Chebyshev polynomials, see Rivlin [49].

Let us consider the Chebyshev series, the orthogonal series based on the Chebyshev polynomials. From (2.55) we see that the Chebyshev coefficients are determined by the ratio of two integrals, and (see Problem 2.2.12) the integral in the denominator is

$$\int_{-1}^{1} (1-x^2)^{-1/2}[T_r(x)]^2 dx = \begin{cases} \pi, & r = 0, \\ \frac{1}{2}\pi, & r > 0. \end{cases} \tag{2.77}$$

Thus the infinite Chebyshev series for f is

$$\frac{1}{2}a_0 + \sum_{r=1}^{\infty} a_r T_r(x), \tag{2.78}$$

where the first coefficient is halved so that the relation

$$a_r = \frac{2}{\pi}\int_{-1}^{1} (1-x^2)^{-1/2} f(x) T_r(x) dx \tag{2.79}$$

holds for all $r \geq 0$. By making the substitution $x = \cos\theta$, we can express (2.79) alternatively in the form

$$a_r = \frac{2}{\pi}\int_0^{\pi} f(\cos\theta)\cos r\theta\, d\theta, \quad r \geq 0. \tag{2.80}$$

We will write

$$f(x) \sim \frac{1}{2}a_0 + \sum_{r=1}^{\infty} a_r T_r(x) \tag{2.81}$$

to signify that the series on the right of (2.81) is the Chebyshev series for the function f. (It should cause no confusion that we used the symbol $\sim$ earlier in a different sense in the statement of Theorem 2.1.8.)

Example 2.2.1 Let us derive the Chebyshev series for $\sin^{-1} x$. On substituting $x = \cos\theta$, we have $\sin^{-1} x = \frac{\pi}{2} - \theta$, and obtain from (2.80) that

$$a_0 = \frac{2}{\pi}\int_0^{\pi} (\tfrac{\pi}{2} - \theta)\, d\theta = 0$$

and

$$a_r = \frac{2}{\pi}\int_0^{\pi} (\tfrac{\pi}{2} - \theta)\cos r\theta\, d\theta = \frac{2}{\pi r}\int_0^{\pi} \sin r\theta\, d\theta, \quad r > 0,$$

on using integration by parts. Thus

$$a_r = \frac{2}{\pi r^2}(1 - (-1)^r), \quad r > 0,$$

which is zero when $r > 0$ is even, and we obtain

$$\sin^{-1} x \sim \frac{4}{\pi}\sum_{r=1}^{\infty} \frac{T_{2r-1}(x)}{(2r-1)^2}. \quad \blacksquare$$

If f is sufficiently differentiable, we can use (2.65) and (2.69) in (2.79) to show that the Chebyshev coefficients for f can be expressed in the form

$$a_r = \frac{2^{r+1}\, r!}{\pi\,(2r)!} \int_{-1}^{1} f^{(r)}(x)(1-x^2)^{r-1/2}dx. \tag{2.82}$$

We see from (2.82) that if f and all its derivatives are nonnegative on $[-1, 1]$, then all its Chebyshev coefficients are nonnegative. It is not hard to see that the same holds for the coefficients of any Jacobi series. We also have the following estimate for the Chebyshev coefficients, which can be justified by using a similar method to that used in proving the analogous result for the Legendre coefficients in Theorem 2.1.6. See Problem 2.2.17.

Theorem 2.2.3 If $f^{(r)}$ is continuous in $(-1, 1)$, then the Chebyshev coefficient a_r is given by

$$a_r = \frac{1}{2^{r-1}\, r!}\, f^{(r)}(\xi_r), \tag{2.83}$$

where $\xi_r \in (-1, 1)$. ■

Example 2.2.2 Consider (2.82) when $f(x) = e^x$, and write

$$I_r = \int_{-1}^{1} e^x (1-x^2)^{r-1/2}dx.$$

Then, on integrating by parts twice, we find that

$$I_r = -(2r-1)(2r-2)I_{r-1} + (2r-1)(2r-3)I_{r-2}, \quad r \geq 2.$$

From (2.82) we see that the Chebyshev coefficient a_r for e^x is

$$a_r = \frac{2^{r+1}\, r!}{\pi\,(2r)!} I_r,$$

and thus we obtain

$$a_r = -(2r-2)a_{r-1} + a_{r-2}, \quad r \geq 2.$$

Like the recurrence relation that we derived in (2.41) for the Legendre coefficients for e^x, this recurrence relation is numerically unstable and is thus of limited practical use. ■

The members of any orthogonal system form a basis for the polynomials. Thus, given an infinite power series, we could transform it to give a series involving the terms of a given orthogonal system. We will illustrate this with the Chebyshev polynomials, which are particularly easy to manipulate. Let us begin by writing

$$x = \cos\theta = \frac{1}{2}\left(e^{i\theta} + e^{-i\theta}\right),$$

where $i^2 = -1$, and we can also write

$$T_n(x) = \cos n\theta = \frac{1}{2}\left(e^{in\theta} + e^{-in\theta}\right). \tag{2.84}$$

Then we have

$$x^n = \frac{1}{2^n}\left(e^{i\theta} + e^{-i\theta}\right)^n,$$

and on using the binomial expansion, we obtain

$$x^n = \frac{1}{2^n}\sum_{r=0}^{n}\binom{n}{r} e^{ir\theta}e^{-i(n-r)\theta} = \frac{1}{2^n}\sum_{r=0}^{n}\binom{n}{r} e^{i(2r-n)\theta}. \tag{2.85}$$

We can combine the rth and $(n-r)$th terms within the latter sum in (2.85), using (2.84) and the fact that the two binomial coefficients involved are equal, to give

$$\binom{n}{r} e^{i(2r-n)\theta} + \binom{n}{n-r} e^{-i(2r-n)\theta} = 2\binom{n}{r} T_{n-2r}(x). \tag{2.86}$$

If n is odd, the terms in the sum in (2.85) combine in pairs, as in (2.86), to give a sum of multiples of odd-order Chebyshev polynomials. We obtain

$$x^{2n+1} = \frac{1}{2^{2n}}\sum_{r=0}^{n}\binom{2n+1}{r} T_{2n+1-2r}(x). \tag{2.87}$$

If n is even in (2.85), we have an odd number of terms in the sum, of which all but one pair off to give the terms involving the even-order Chebyshev polynomials $T_n, T_{n-2}, \ldots, T_2$, leaving a single term, which corresponds to $r = \frac{1}{2}n$ and gives the contribution involving T_0. Thus we obtain

$$x^{2n} = \frac{1}{2^{2n}}\binom{2n}{n} T_0(x) + \frac{1}{2^{2n-1}}\sum_{r=0}^{n-1}\binom{2n}{r} T_{2n-2r}(x). \tag{2.88}$$

Now suppose we have a function f expressed as an infinite power series,

$$f(x) = \sum_{r=0}^{\infty} c_r x^r, \tag{2.89}$$

and let this series be uniformly convergent on $[-1, 1]$. We then express each monomial x^r in terms of the Chebyshev polynomials, using (2.87) when r is of the form $2n+1$, and (2.88) when r is of the form $2n$. On collecting together all terms involving T_0, T_1, and so on, we obtain a series of the form

$$f(x) = \frac{1}{2}a_0 + \sum_{r=1}^{\infty} a_r T_r(x). \tag{2.90}$$

It follows from the uniform convergence of the power series in (2.89) that the series defined by (2.90) is also uniformly convergent. Thus, for any integer $s \geq 0$, we may multiply (2.90) throughout by $(1-x^2)^{-1/2}T_s(x)$ and integrate term by term over $[-1,1]$. Due to the orthogonality property (2.66), we obtain

$$\int_{-1}^{1}(1-x^2)^{-1/2}f(x)T_s(x)dx = a_s\int_{-1}^{1}(1-x^2)^{-1/2}[T_s(x)]^2dx, \quad s>0,$$

and when $s=0$ we have

$$\int_{-1}^{1}(1-x^2)^{-1/2}f(x)dx = \frac{1}{2}a_0\int_{-1}^{1}(1-x^2)^{-1/2}dx.$$

In view of (2.77) and (2.79), we can see that (2.90) is *the* Chebyshev series for f. The relation between the Chebyshev coefficients a_r and the coefficients c_r in the power series then follows from (2.87) and (2.88). We obtain

$$a_r = \frac{1}{2^{r-1}}\sum_{s=0}^{\infty}\frac{1}{2^{2s}}\binom{r+2s}{s}c_{r+2s}, \tag{2.91}$$

for all $r \geq 0$.

Example 2.2.3 With $f(x)=e^x$ in (2.89), we have $c_r = 1/r!$, and we see from (2.91) that the Chebyshev coefficients for e^x are

$$a_r = \frac{1}{2^{r-1}}\sum_{s=0}^{\infty}\frac{1}{2^{2s}}\cdot\frac{1}{s!(r+s)!}, \tag{2.92}$$

for all $r \geq 0$. The first two Chebyshev coefficients for e^x, rounded to six decimal places, are $a_0 = 2.532132$ and $a_1 = 1.130318$. The next few are given in the following table:

r	2	3	4	5	6	7
a_r	0.271495	0.044337	0.005474	0.000543	0.000045	0.000003

We note from (2.92) that for large r,

$$a_r \sim \frac{1}{2^{r-1}}\cdot\frac{1}{r!},$$

compared with $c_r = 1/r!$. ■

The inner product (2.42) can be generalized by incorporating a weight function ω, writing

$$(f,g) = \int_{-1}^{1}\omega(x)f(x)g(x)dx. \tag{2.93}$$

Then we can show that

$$(f - p_n, q_s^\omega) = 0, \quad 0 \le s \le n, \tag{2.94}$$

where p_n is the partial orthogonal series for f with respect to the weight function ω, and the polynomials q_s^ω are orthogonal on $[-1, 1]$ with respect to ω. We may justify (2.94) in the same way that we justified its special case (2.44). Then, using the same method of proof as we used for Theorem 2.1.7, we obtain the following generalization.

Theorem 2.2.4 The error term $f - p_n$ changes sign on at least $n+1$ points in the interior of $[-1, 1]$, where p_n is the partial sum of the orthogonal series for f with respect to a given weight function ω. ■

In particular, the error term of the partial Chebyshev series $f - p_n$ changes sign on at least $n+1$ points in the interior of $[-1, 1]$. We now consider three theorems that generalize results, stated earlier, related to the Legendre polynomials.

Theorem 2.2.5 Given a weight function ω, the polynomial $p \in P_n$ that minimizes

$$\int_{-1}^{1} \omega(x)[x^{n+1} - p(x)]^2 dx$$

is $p(x) = x^{n+1} - q_{n+1}^\omega(x)$, where q_{n+1}^ω is the orthogonal polynomial of degree $n + 1$, with leading coefficient unity, with respect to the weight function ω.

Proof. The proof is similar to that used in Problem 2.1.8 to verify the special case where $\omega(x) = 1$. ■

Theorem 2.2.6 Given the norm (2.53) based on the weight function ω, the minimum value of $\|(x - x_0) \cdots (x - x_n)\|$ is attained when the x_j are the zeros of the orthogonal polynomial q_{n+1}^ω.

Proof. This result follows immediately from Theorem 2.2.5. ■

Theorem 2.2.7 Given the norm (2.53) based on the weight function ω, let p_n denote the best weighted square norm approximation, with respect to ω, for a given function f. Then if $f^{(n+1)}$ is continuous on $[-1, 1]$, there exists a number ζ in $(-1, 1)$ such that

$$\|f - p_n\| = \frac{|f^{(n+1)}(\zeta)|}{(n+1)!} \|(x - x_0^*) \cdots (x - x_n^*)\|, \tag{2.95}$$

where the x_j^* denote the zeros of the orthogonal polynomial q_{n+1}^ω.

Proof. This result may be justified in the same way as Theorem 2.1.8, which is concerned with the special case where $\omega(x) = 1$. ■

A special case of the last theorem is Theorem 2.1.8 concerning the error of the truncated Legendre series. Another special case of Theorem 2.2.7, which we now give as a separate theorem, concerns the error of the Chebyshev series.

Theorem 2.2.8 Let p_n denote the truncated Chebyshev series of degree n for f on $[-1, 1]$. Then if $f^{(n+1)}$ is continuous on $[-1, 1]$, there exists a number ζ in $(-1, 1)$ such that

$$\|f - p_n\| = \frac{\sqrt{\pi}}{2^{n+1/2}} \frac{|f^{(n+1)}(\zeta)|}{(n+1)!}, \tag{2.96}$$

where the norm is given by (2.53) with $\omega(x) = (1 - x^2)^{-1/2}$.

Proof. This result follows from Theorem 2.2.7 and Problem 2.2.18. ■

We continue this section with a brief account of the system of polynomials that are orthogonal on $[-1, 1]$ with respect to the weight function $(1-x^2)^{1/2}$. As we already mentioned, these are called the Chebyshev polynomials of the second kind, a special case of the ultraspherical polynomials. Like *the* Chebyshev polynomials T_n, the Chebyshev polynomials of the second kind can be expressed simply in terms of circular functions. For $n \geq 0$, let us write

$$U_n(x) = \frac{\sin(n+1)\theta}{\sin\theta}, \quad \text{where } x = \cos\theta, \quad -1 \leq x \leq 1. \tag{2.97}$$

On making the substitution $x = \cos\theta$, the interval $-1 \leq x \leq 1$ is mapped to $0 \leq \theta \leq \pi$, and we obtain

$$\int_{-1}^{1} (1 - x^2)^{1/2} U_r(x) U_s(x) dx = \int_0^{\pi} \sin(r+1)\theta \sin(s+1)\theta \, d\theta, \quad r \neq s.$$

Since

$$\sin(r+1)\theta \sin(s+1)\theta = \frac{1}{2}(\cos(r-s)\theta - \cos(r+s+2)\theta),$$

we find that

$$\int_{-1}^{1} (1 - x^2)^{1/2} U_r(x) U_s(x) dx = 0, \quad r \neq s,$$

showing that the functions U_r are orthogonal on $[-1, 1]$ with respect to the weight function $(1 - x^2)^{1/2}$. It remains to show that $U_n \in P_n$. Now,

$$\sin(n+2)\theta + \sin n\theta = 2\sin(n+1)\theta\cos\theta,$$

and if we divide throughout by $\sin\theta$ and use (2.97), we obtain the recurrence relation

$$U_{n+1}(x) = 2xU_n(x) - U_{n-1}(x), \quad n \geq 1. \tag{2.98}$$

We observe from (2.97) that $U_0(x) = 1$ and $U_1(x) = 2x$, and it follows from (2.98) by induction that $U_n \in P_n$ for each n. Thus (2.97) defines a system of polynomials that are orthogonal on $[-1, 1]$ with respect to the weight function $(1 - x^2)^{1/2}$.

Since $x = 1$ corresponds to $\theta = 0$ under the transformation $x = \cos\theta$, for $x \in [-1, 1]$, we have

$$U_n(1) = \lim_{\theta \to 0} \frac{\sin(n+1)\theta}{\sin\theta} = n + 1, \quad n \geq 0,$$

using L'Hospital's rule. Since, as for all the ultraspherical polynomials, U_n is even or odd when n is even or odd, $U_n(-1) = (-1)^n(n + 1)$. The zeros of U_n (see Problem 2.2.21) are

$$x = \cos(j\pi/(n+1)), \quad 1 \leq j \leq n.$$

Let us consider the orthogonal series based on the Chebyshev polynomials of the second kind, whose weight function is $(1 - x^2)^{1/2}$. As we saw in (2.55), an orthogonal coefficient is determined by the ratio of two integrals. Let us begin with the integral in the denominator on the right side of (2.55), with $\omega(x) = (1 - x^2)^{1/2}$, and make the substitution $x = \cos\theta$ to give

$$\int_{-1}^{1} (1 - x^2)^{1/2} [U_n(x)]^2 dx = \int_0^{\pi} \sin^2(n+1)\theta \, d\theta = \frac{1}{2}\pi,$$

for all $n \geq 0$. Thus if b_n denotes the coefficient of U_n in the orthogonal series, it follows from (2.55) that

$$b_n = \frac{2}{\pi} \int_0^{\pi} \sin\theta \sin(n+1)\theta \, f(\cos\theta) d\theta, \tag{2.99}$$

for all $n \geq 0$.

We can derive a simple relation between the coefficients b_n of a series of the second kind, defined by (2.99), and the coefficients a_n of *the* Chebyshev series, defined by (2.80). For we can write

$$2\sin\theta \sin(n+1)\theta = \cos n\theta - \cos(n+2)\theta,$$

for all $n \geq 0$, and thus, comparing (2.80) and (2.99), we have

$$b_n = \frac{1}{2}(a_n - a_{n+2}), \quad n \geq 0. \tag{2.100}$$

We can obtain the relation (2.100) otherwise by formally comparing the first and second kinds of Chebyshev expansions of a given function, as we did above in deriving (2.91) by comparing a Chebyshev series with a power series. We begin by writing

$$\sin(n+1)\theta - \sin(n-1)\theta = 2\cos n\theta \sin\theta.$$

On dividing throughout by $\sin\theta$, we obtain

$$U_n(x) - U_{n-2}(x) = 2T_n(x), \quad n \geq 2, \tag{2.101}$$

and we also have

$$U_0(x) = T_0(x) \qquad \text{and} \qquad U_1(x) = 2T_1(x). \tag{2.102}$$

Then

$$\frac{1}{2}a_0T_0(x) + \sum_{r=1}^{\infty} a_rT_r(x) = \frac{1}{2}\sum_{r=0}^{1} a_rU_r(x) + \frac{1}{2}\sum_{r=2}^{\infty} a_r(U_r(x) - U_{r-2}(x)).$$

If we now express the right side of the latter equation as

$$\frac{1}{2}\sum_{r=0}^{1} a_rU_r(x) + \frac{1}{2}\sum_{r=2}^{\infty} a_r(U_r(x) - U_{r-2}(x)) = \sum_{r=0}^{\infty} b_rU_r(x),$$

it is clear that the relation between the b_n and the a_n is as given above in (2.100).

To complete this section on the Chebyshev polynomials, we will consider again the Hermite interpolating polynomial $p_{2n+1}(x)$, defined in (1.38), which interpolates a given function $f(x)$, and whose first derivative interpolates $f'(x)$, at $n+1$ arbitrary abscissas $x_0, x_1, \ldots, x_n$. Let us take the x_i to be the zeros of T_{n+1} arranged in the order $-1 < x_0 < x_1 < \cdots < x_n < 1$, so that

$$x_i = \cos\frac{(2n-2i+1)\pi}{2n+2}, \quad 0 \leq i \leq n.$$

It then follows from Problem 1.1.3 that the fundamental polynomial on the zeros of T_{n+1} is

$$L_i(x) = \frac{T_{n+1}(x)}{(x-x_i)T'_{n+1}(x_i)},$$

and on using the expression for the derivative of the Chebyshev polynomial in Problem 2.2.11, we can write this as

$$L_i(x) = (-1)^{n-i}\frac{T_{n+1}(x)\sqrt{1-x_i^2}}{(n+1)(x-x_i)}.$$

It also follows from Problem 1.1.3 and the expressions for the first two derivatives of the Chebyshev polynomial in Problem 2.2.11 that

$$L'_i(x_i) = \frac{1}{2}\frac{T''_{n+1}(x_i)}{T'_{n+1}(x_i)} = \frac{1}{2}\frac{x_i}{1-x_i^2}.$$

On substituting these results into (1.39) and (1.40), we find that the Hermite interpolating polynomial on the zeros of T_{n+1} is given by

$$p_{2n+1}(x) = \sum_{i=0}^{n} [f(x_i)u_i(x) + f'(x_i)v_i(x)],$$

where

$$u_i(x) = \left(\frac{T_{n+1}(x)}{n+1} \right)^2 \frac{1 - x_i x}{(x - x_i)^2} \tag{2.103}$$

and

$$v_i(x) = \left(\frac{T_{n+1}(x)}{n+1} \right)^2 \left(\frac{1 - x_i^2}{x - x_i} \right). \tag{2.104}$$

Problem 2.2.1 Use integration by parts on (2.62) to show that

$$\Gamma(x+1) = x\Gamma(x).$$

Show that $\Gamma(1) = 1$, and $\Gamma(n+1) = n!$, for all integers $n \geq 0$, and thus verify (2.61).

Problem 2.2.2 Use the Leibniz rule (1.83) to obtain

$$Q_n^{(\alpha,\beta)}(-1) = (-1)^n \binom{n+\beta}{n} = (-1)^n \frac{\Gamma(n+\beta+1)}{\Gamma(n+1)\Gamma(\beta+1)},$$

a companion formula for (2.61).

Problem 2.2.3 Verify that Theorem 2.2.1 reduces to Theorem 2.1.4 when we put $\alpha = \beta = 0$.

Problem 2.2.4 Deduce from (2.60) that

$$Q_n^{(-1/2,-1/2)}(1) = \frac{1}{2^{2n}} \binom{2n}{n},$$

and use Stirling's formula (see Problem 2.1.12) to show that

$$Q_n^{(-1/2,-1/2)}(1) \sim \frac{1}{\sqrt{\pi n}}.$$

Problem 2.2.5 Verify from Theorem 2.2.1 that

$$Q_{n+1}^{(-1/2,-1/2)}(x) = a_n x Q_n^{(-1/2,-1/2)}(x) - c_n Q_{n-1}^{(-1/2,-1/2)}(x),$$

where

$$a_n = \frac{2n+1}{n+1} \quad \text{and} \quad c_n = \frac{4n^2-1}{4n(n+1)},$$

with $Q_0^{(-1/2,-1/2)}(x) = 1$ and $Q_1^{(-1/2,-1/2)}(x) = \frac{1}{2}x$.

Problem 2.2.6 For $x \geq 1$, write

$$x = \cosh\theta = \frac{1}{2}(e^{\theta} + e^{-\theta}).$$

Verify that $\cosh\theta \geq 1$ for all real θ and show also that

$$\cosh(n+1)\theta + \cosh(n-1)\theta = 2\cosh\theta\cosh n\theta.$$

Hence verify by induction, using the recurrence relation (2.72), that

$$T_n(x) = \cosh n\theta, \quad \text{where} \quad x = \cosh\theta,$$

for $x \geq 1$. Finally, show that

$$T_n(x) = (-1)^n \cosh n\theta, \quad \text{where} \quad x = -\cosh\theta,$$

for $x \leq -1$.

Problem 2.2.7 Deduce from the recurrence relation (2.72) that $T_n(1) = 1$ and $T_n(-1) = (-1)^n$ for all $n \geq 0$.

Problem 2.2.8 Using the recurrence relation (2.72), verify that

$$T_5(x) = 16x^5 - 20x^3 + 5x \quad \text{and} \quad T_6(x) = 32x^6 - 48x^4 + 18x^2 - 1.$$

Write $T_2(T_3(x)) = 2(T_3(x))^2 - 1$, and show directly that this simplifies to give $T_6(x)$. Show also that $T_3(T_2(x)) = T_6(x)$. More generally, show that

$$T_m(T_n(x)) = T_n(T_m(x)) = T_{mn}(x),$$

for all integers $m, n \geq 0$.

Problem 2.2.9 Verify that

$$T_{2n+2}(x) = 2T_2(x)T_{2n}(x) - T_{2n-2}(x),$$

for $n \geq 1$, with $T_0(x) = 1$ and $T_2(x) = 2x^2 - 1$, giving a recurrence relation that computes only the even-order Chebyshev polynomials. Find a recurrence relation that computes only the odd-order Chebyshev polynomials.

Problem 2.2.10 Verify that

$$T_{n+k}(x)T_{n-k}(x) - (T_n(x))^2 = (T_k(x))^2 - 1$$

for all $n \geq k \geq 0$.

Problem 2.2.11 Use the chain rule of differentiation to show from the definition of the Chebyshev polynomials in (2.71) that

$$T_n'(x) = \frac{n\sin n\theta}{\sin\theta} \quad \text{and} \quad T_n''(x) = \frac{-n^2\sin\theta\cos n\theta + n\sin n\theta\cos\theta}{\sin^3\theta},$$

where $x = \cos\theta$, and deduce that T_n satisfies the second-order differential equation

$$(1-x^2)T_n''(x) - xT_n'(x) + n^2T_n(x).$$

Problem 2.2.12 By making the substitution $x = \cos\theta$, show that

$$\int_{-1}^{1}(1-x^2)^{-1/2}[T_r(x)]^2 dx = \int_0^{\pi}\cos^2 r\theta\, d\theta = \frac{1}{2}\int_0^{\pi}(1+\cos 2r\theta)\, d\theta,$$

and hence show that

$$\int_{-1}^{1}(1-x^2)^{-1/2}[T_r(x)]^2 dx = \begin{cases} \pi, & r = 0, \\ \frac{1}{2}\pi, & r > 0. \end{cases}$$

Problem 2.2.13 Derive the Chebyshev series for $[(1+x)/2]^{1/2}$. Hint: Make the substitution $x = \cos\theta$ and use the identity

$$2\cos\frac{1}{2}\theta\cos r\theta = \cos\left(r+\frac{1}{2}\right)\theta + \cos\left(r-\frac{1}{2}\right)\theta.$$

Problem 2.2.14 Obtain the Chebyshev series for $\cos^{-1} x$.

Problem 2.2.15 Find the Chebyshev series for $(1-x^2)^{1/2}$.

Problem 2.2.16 Assuming that the Chebyshev series for $\sin^{-1} x$, derived in Example 2.2.1, converges uniformly to $\sin^{-1} x$ on $[-1, 1]$, deduce that

$$\frac{1}{1^2} + \frac{1}{3^2} + \frac{1}{5^2} + \frac{1}{7^2} + \cdots = \frac{\pi^2}{8}.$$

Problem 2.2.17 Let

$$J_r = \int_{-1}^{1}(1-x^2)^{r-1/2}dx.$$

Verify that $J_0 = \pi$, as obtained in Problem 2.2.12, and use integration by parts to show that

$$J_r = (2r-1)\int_{-1}^{1}x^2(1-x^2)^{r-3/2}dx = (2r-1)(J_{r-1} - J_r).$$

Deduce that

$$J_r = \frac{\pi\,(2r)!}{2^{2r}\,(r!)^2},$$

and hence verify (2.83).

Problem 2.2.18 Let $x_0^*, \ldots, x_n^*$ denote the zeros of the Chebyshev polynomial T_{n+1}, and let $\|\cdot\|$ denote the weighted square norm, defined by (2.53), with weight function $(1-x^2)^{-1/2}$. Use the substitution $x = \cos\theta$ to show that

$$\|(x-x_0^*)\cdots(x-x_n^*)\| = \frac{1}{2^n}\left(\int_{-1}^{1}(1-x^2)^{-1/2}[T_{n+1}(x)]^2 dx\right)^{1/2} = \frac{\pi^{1/2}}{2^{n+1/2}}.$$

Problem 2.2.19 Begin with the Maclaurin series

$$\cos \tfrac{1}{2}\pi x = \sum_{r=0}^{\infty}(-1)^r \frac{(\pi x)^{2r}}{2^{2r}(2r)!},$$

and use the relation (2.91) to compute the first few Chebyshev coefficients for $\cos(\frac{1}{2}\pi x)$.

Problem 2.2.20 In the text it is shown that

$$U_n(1) = n+1 \qquad \text{and} \qquad U_n(-1) = (-1)^n(n+1).$$

Deduce these results alternatively by using an induction argument on the recurrence relation for U_n.

Problem 2.2.21 Deduce from (2.97) that the Chebyshev polynomial of the second kind U_n is zero for values of θ such that $\sin(n+1)\theta = 0$ and $\sin\theta$ is nonzero. Show that

$$\sin(n+1)\theta = 0 \quad \Rightarrow \quad (n+1)\theta = j\pi \quad \Rightarrow \quad x = \cos(j\pi/(n+1)),$$

and hence show that U_n has all its n zeros in $(-1,1)$, at the abscissas $x = \cos(j\pi/(n+1))$, for $1 \le j \le n$.

Problem 2.2.22 Use the relations (2.101) and (2.102), which express the Chebyshev polynomials T_n, to derive an expression for every monomial in terms of the U_n, like those given for x^{2n+1} and x^{2n} in terms of the T_n, in (2.87) and (2.88), respectively.

Problem 2.2.23 Deduce from (2.71) and (2.97) that

$$T'_{n+1}(x) = (n+1)U_n(x),$$

and hence show that $T'_n(1) = n^2$.

Problem 2.2.24 Show that

$$U_n(x) = T_n(x) + xU_{n-1}(x), \quad n \ge 1,$$

and deduce that

$$U_n(x) = \sum_{r=0}^{n} x^r T_{n-r}(x), \quad n \ge 0.$$

Problem 2.2.25 Deduce from (2.101) that

$$U_{2n+1}(x) = 2\sum_{r=0}^{n} T_{2n+1-2r}(x)$$

and

$$U_{2n}(x) = T_0(x) + 2\sum_{r=0}^{n-1} T_{2n-2r}(x).$$

2.3 Finite Point Sets

In the last section we constructed a sequence of polynomials (q_n^ω) that are orthogonal on $[-1,1]$ with respect to a given weight function ω, and we used these as a basis for constructing best approximations for functions defined on $[-1,1]$.

We will now consider best approximations for functions that are defined on a finite set of points, say $X = \{x_0, x_1, \ldots, x_N\}$. Associated with each point x_j we will assign a positive number ω_j, called a *weight*. Then, given a function f defined on the point set X, we seek a polynomial $p \in P_n$ that minimizes

$$\sum_{j=0}^{N} \omega_j [f(x_j) - p(x_j)]^2, \tag{2.105}$$

which we will call a least squares approximation to f on X with respect to the given weights. We now define

$$\|g\| = \left(\sum_{j=0}^{N} \omega_j [g(x_j)]^2 \right)^{1/2}, \tag{2.106}$$

and we can show that this is a norm, as in Definition 2.1.1. It is analogous to the norm defined by (2.53).

Following the same method as we used in Section 2.2 for weighted square norm approximations on the interval $[-1,1]$, we construct a sequence of polynomials $(q_n^\omega)_{n=0}^N$, where q_n^ω is of degree n, has leading coefficient unity, and satisfies

$$\sum_{j=0}^{N} \omega_j x_j^r q_n^\omega(x_j) = 0, \qquad 0 \le r < n. \tag{2.107}$$

We find that the orthogonal polynomials q_n^ω satisfy the recurrence relation

$$q_{n+1}^\omega(x) = (x - \alpha_n) q_n^\omega(x) - \beta_n q_{n-1}^\omega(x), \tag{2.108}$$

where

$$\alpha_n = \sum_{j=0}^{N} \omega_j x_j [q_n^\omega(x_j)]^2 \Big/ \sum_{j=0}^{N} \omega_j [q_n^\omega(x_j)]^2 \tag{2.109}$$

and

$$\beta_n = \sum_{j=0}^{N} \omega_j [q_n^\omega(x_j)]^2 \Big/ \sum_{j=0}^{N} \omega_j [q_{n-1}^\omega(x_j)]^2. \tag{2.110}$$

Note that the last two relations are analogous to (2.51) and (2.52), respectively. We then discover that

$$\|f - p\| = \left(\sum_{j=0}^{N} \omega_j [f(x_j) - p(x_j)]^2 \right)^{1/2}$$

is minimized over all $p \in P_n$, with $n \leq N$, by choosing $p = p_n$, where

$$p_n(x) = \sum_{r=0}^{n} a_r q_r^{\omega}(x) \tag{2.111}$$

and

$$a_r = \sum_{j=0}^{N} \omega_j f(x_j) q_r^{\omega}(x_j) \Big/ \sum_{j=0}^{N} \omega_j [q_r^{\omega}(x_j)]^2, \quad 0 \leq r \leq n. \tag{2.112}$$

We note that p_N must be the interpolating polynomial for f on the set X, since then $\|f - p_N\| = 0$, and this shows why we have restricted n to be not greater than N. We also observe that the sequence of orthogonal polynomials is unchanged if we multiply all the weights by any positive constant.

It can sometimes be more convenient to work with orthogonal polynomials that do not have leading coefficient 1. Suppose $Q_r^{\omega}(x) = c_r q_r^{\omega}(x)$, where $c_r \neq 0$ may depend on r, but is independent of x. Then it follows from (2.111) and (2.112) that if we express the best square norm approximation to f on X as

$$p_n(x) = \sum_{r=0}^{n} a_r Q_r^{\omega}(x), \tag{2.113}$$

with $n \leq N$, then the coefficient a_r is given by

$$a_r = \sum_{j=0}^{N} \omega_j f(x_j) Q_r^{\omega}(x_j) \Big/ \sum_{j=0}^{N} \omega_j [Q_r^{\omega}(x_j)]^2, \quad 0 \leq r \leq n. \tag{2.114}$$

Note that the polynomials Q_r^{ω} will satisfy the simple recurrence relation (2.108) only if every c_r is equal to 1.

Example 2.3.1 To illustrate the foregoing material, let us choose the set X as $\{-1, 0, 1\}$, with weights ω_j all equal, and let f be defined on X by the following table:

x	-1	0	1
$f(x)$	1	2	4

We have $q_0^{\omega}(x) = 1$ and $q_1^{\omega}(x) = x$. Then we find that $\alpha_1 = 0$ and $\beta_1 = \frac{2}{3}$, so that $q_2^{\omega}(x) = x^2 - \frac{2}{3}$. On using (2.112) we find that the orthogonal coefficients are $a_0 = \frac{7}{3}$, $a_1 = \frac{3}{2}$, and $a_2 = \frac{1}{2}$. Thus the best approximation in P_1 is $\frac{7}{3} + \frac{3}{2}x$, while that in P_2 is $\frac{7}{3} + \frac{3}{2}x + \frac{1}{2}(x^2 - \frac{2}{3})$. As expected, the last polynomial interpolates f on X. ■

We require some further notation, writing $\sum{}'$ to denote a sum in which the first term is halved, and $\sum{}''$ to denote a sum in which both the first term and the last terms are halved. Thus

$$\sum_{j=0}^{N}{}' u_j = \frac{1}{2}u_0 + u_1 + \cdots + u_N \tag{2.115}$$

and

$$\sum_{j=0}^{N}{}'' u_j = \frac{1}{2}u_0 + u_1 + \cdots + u_{N-1} + \frac{1}{2}u_N. \tag{2.116}$$

Now let $X = \{x_0, x_1, \ldots, x_N\}$, where $x_j = \cos(\pi j/N),\ 0 \leq j \leq N$, and let

$$\omega_j = \begin{cases} \frac{1}{2}, & j = 0 \text{ and } N, \\ 1, & 1 \leq j \leq N-1. \end{cases}$$

Thus X is the set of extreme points of the Chebyshev polynomial T_N. We can verify (see Problem 2.3.1) that

$$\sum_{j=0}^{N}{}'' T_r(x_j)T_s(x_j) = 0, \qquad r \neq s, \tag{2.117}$$

and so the Chebyshev polynomials are orthogonal on the extreme points of T_N with respect to the given set of weights. It then follows from (2.113) and (2.114) that the best weighted square norm approximation with respect to this set of points and weights is

$$p_n(x) = \sum_{r=0}^{n} a_r T_r(x), \tag{2.118}$$

where

$$a_r = \sum_{j=0}^{N}{}'' f(x_j)T_r(x_j) \Big/ \sum_{j=0}^{N}{}'' [T_r(x_j)]^2, \quad 0 \leq r \leq n. \tag{2.119}$$

We find (again see Problem 2.3.1) that

$$\sum_{j=0}^{N}{}'' [T_r(x_j)]^2 = \begin{cases} N, & r = 0 \text{ or } N, \\ \frac{1}{2}N, & 1 \leq r \leq N-1. \end{cases} \tag{2.120}$$

On combining (2.118), (2.119), and (2.120) we see that the best weighted square norm approximation to f on the extreme points of T_N is

$$p_n(x) = \sum_{r=0}^{n}{}' \alpha_r T_r(x), \quad n \leq N, \tag{2.121}$$

say, where

$$\alpha_r = \frac{2}{N} \sum_{j=0}^{N}{}'' f(x_j) T_r(x_j), \quad 0 \le r \le n. \tag{2.122}$$

If we choose $n = N$ in (2.121), the least squares approximation must coincide with the interpolating polynomial for f on the extreme points of T_N. We now derive a connection between the above coefficients α_r and the Chebyshev coefficients a_r, defined by (2.79). First we note that

$$T_{2kN\pm r}(x_j) = \cos\left(\frac{(2kN \pm r)\pi j}{N}\right) = \cos\left(2k\pi j \pm \frac{r\pi j}{N}\right) = \cos\left(\frac{r\pi j}{N}\right),$$

and thus

$$T_{2kN\pm r}(x_j) = T_r(x_j). \tag{2.123}$$

Let us assume that *the* Chebyshev series for f, given by (2.81), converges uniformly to f on $[-1, 1]$. Then it follows from (2.122) that

$$\begin{aligned} \alpha_r &= \frac{2}{N} \sum_{j=0}^{N}{}'' \left(\sum_{s=0}^{\infty}{}' a_s T_s(x_j) \right) T_r(x_j) \\ &= \frac{2}{N} \sum_{s=0}^{\infty}{}' a_s \sum_{j=0}^{N}{}'' T_s(x_j) T_r(x_j). \end{aligned} \tag{2.124}$$

We observe from (2.123) and the orthogonality property (2.117) that the only nonzero summations over j in the second line of (2.124) are those for which $s = r$, $2N - r$, $2N + r$, $4N - r$, $4N + r$, and so on. Thus, for $r \ne 0$ or N,

$$\alpha_r = a_r + \sum_{k=1}^{\infty} (a_{2kN-r} + a_{2kN+r}), \tag{2.125}$$

and we find that (2.125) holds also for $r = 0$ and N, on examining these cases separately.

We saw in the last section that the Chebyshev polynomials are orthogonal on the interval $[-1, 1]$ with respect to a certain weight function, and we have seen above that they are orthogonal on the extreme points of T_N with respect to certain weights. We now show that they satisfy a third orthogonality property. For the Chebyshev polynomials are also orthogonal on the set $X = \{x_1^*, \ldots, x_N^*\}$ with all weights equal, where the x_j^* are the zeros of T_N. We can verify that (see Problem 2.3.2)

$$\sum_{j=1}^{N} T_r(x_j^*) T_s(x_j^*) = \begin{cases} 0, & r \ne s \text{ or } r = s = N, \\ \frac{1}{2}N, & r = s \ne 0 \text{ or } N, \\ N, & r = s = 0. \end{cases} \tag{2.126}$$

Then, following the same method as we used in deriving (2.121) and (2.122), we find that the best square norm approximation to the function f on the zeros of T_N is

$$p_n(x) = \sum_{r=0}^{n}{}' \alpha_r^* T_r(x), \qquad n \le N-1, \tag{2.127}$$

where

$$\alpha_r^* = \frac{2}{N} \sum_{j=1}^{N} f(x_j^*) T_r(x_j^*). \tag{2.128}$$

If we choose $n = N-1$ in (2.127), the least squares approximation is just the interpolating polynomial for f on the zeros of T_N.

We can express α_r^* as a sum involving the Chebyshev coefficients, using the same method as we used above to derive the expression (2.125) for α_r. We first verify that

$$T_{2kN\pm r}(x_j^*) = (-1)^k T_r(x_j^*). \tag{2.129}$$

Assuming that the Chebyshev series for f converges uniformly to f on $[-1,1]$, we find that

$$\alpha_r^* = a_r + \sum_{k=1}^{\infty} (-1)^k (a_{2kN-r} + a_{2kN+r}) \tag{2.130}$$

for $0 \le r \le N-1$, and we see from (2.125) and (2.130) that

$$\frac{1}{2}(\alpha_r + \alpha_r^*) = a_r + \sum_{k=1}^{\infty} (a_{4kN-r} + a_{4kN+r}). \tag{2.131}$$

The three expressions (2.125), (2.130), and (2.131), connecting the a_r with the coefficients α_r and α_r^*, suggest a practical method of computing the Chebyshev coefficients a_r. We choose a value of N and then compute α_r and α_r^*, using (2.125) and (2.130). If α_r and α_r^* differ by more than an acceptable amount, we increase N and recompute α_r and α_r^*. When they agree sufficiently, we use $\frac{1}{2}(\alpha_r + \alpha_r^*)$ as an approximation to a_r.

Problem 2.3.1 Show that $T_r(x_j) = \cos(\pi r j/N)$, where the x_j are the extreme points of T_N, and hence express the left side of (2.117) as a sum of cosines, using the relation

$$\cos(\theta + \phi) + \cos(\theta - \phi) = 2\cos\theta\cos\phi.$$

Evaluate the sum by using the identity

$$\cos k\theta = \frac{\sin(k+\frac{1}{2})\theta - \sin(k-\frac{1}{2})\theta}{2\sin\frac{1}{2}\theta},$$

and hence verify the orthogonality relation (2.117). Similarly verify (2.120).

Problem 2.3.2 With $x_j^* = \cos \frac{(2j-1)\pi}{2N}$, show that

$$T_r(x_j^*) = \cos \frac{(2j-1)r\pi}{2N}$$

and follow the method of the last problem to verify (2.126). Show also that

$$T_{2kN\pm r}(x_j^*) = (-1)^k T_r(x_j^*).$$

Problem 2.3.3 Verify (2.130) and (2.131).

Problem 2.3.4 Show that α_r, defined by (2.122), is the approximation to the Chebyshev coefficient a_r that is obtained by estimating the integral

$$a_r = \frac{2}{\pi} \int_0^\pi f(\cos\theta) \cos r\theta \, d\theta$$

(see (2.80)) using the composite form of the trapezoidal rule. Show also that α_r^*, defined by (2.128), is the approximation obtained by applying the composite midpoint rule to the same integral. (These integration rules are discussed in Chapter 3.)

2.4 Minimax Approximation

A best approximation with respect to the maximum norm, defined by (2.5), is called a *minimax approximation*, since we wish to *minimize*, over all p_n in P_n, the *maximum* value of the error $|f(x) - p_n(x)|$ over $[-1, 1]$. For convenience, we will continue to work with the interval $[-1, 1]$, although what we have to say about minimax approximations may be applied to any finite interval by making a linear change of variable. Minimax approximations, which are also called *uniform approximations*, are important because of the following famous theorem due to Karl Weierstrass (1815–1897).

Theorem 2.4.1 Given any $f \in C[-1, 1]$ and any $\epsilon > 0$, there exists a polynomial p such that

$$|f(x) - p(x)| < \epsilon, \qquad -1 \le x \le 1. \qquad \blacksquare \tag{2.132}$$

There are many proofs of this key theorem and we will give two in Chapter 7, based on proofs given by two mathematicians who were born in the same year, S. N. Bernstein (1880–1968) and L. Fejér (1880–1959). As we will see in this section, the property that characterizes the error $f - p_n$ of a minimax polynomial $p_n \in P_n$ for a function $f \in C[-1, 1]$ is one that is shared with the Chebyshev polynomial T_{n+1}. As we saw in the last section, the polynomial T_{n+1} attains its maximum modulus on $[-1, 1]$ at $n + 2$ points belonging to $[-1, 1]$, and $T_{n+1}(x)$ alternates in sign on these points. It is useful to have a name for this property, which we give now.

Definition 2.4.1 A function $E(x)$ is said to *equioscillate* on $n+2$ points of $[-1,1]$ if for $-1 \leq x_1 < x_2 < \cdots < x_{n+2} \leq 1$,

$$|E(x_j| = \max_{-1 \leq x \leq 1} |E(x)|, \quad 1 \leq j \leq n+2,$$

and

$$E(x_{j+1}) = -E(x_j), \quad 1 \leq j \leq n+1. \qquad \blacksquare$$

Thus T_{n+1} equioscillates on the $n+2$ points $t_j = \cos[(j-1)\pi/(n+1)]$, $1 \leq j \leq n+2$, since $T_{n+1}(t_j) = (-1)^{j-1}$ and the maximum modulus of T_{n+1} on $[-1,1]$ is 1.

Example 2.4.1 Consider the function

$$E(x) = |x| - \frac{1}{8} - x^2$$

on $[-1,1]$. We can easily check that

$$E(-1) = -E(-\tfrac{1}{2}) = E(0) = -E(\tfrac{1}{2}) = E(1).$$

The function E is differentiable on $[-1,1]$ except at $x=0$. For $0 < x \leq 1$, we have $E(x) = x - \frac{1}{8} - x^2$ and

$$E'(x) = 1 - 2x = 0 \;\Rightarrow\; \quad x = \frac{1}{2}.$$

Thus E has a turning value at $x = \frac{1}{2}$, and we similarly find that E has another turning value at $x = -\frac{1}{2}$. It follows that E attains its maximum modulus on $[-1,1]$ at the five points $x = 0, \pm\frac{1}{2}, \pm 1$, and E equioscillates on these points. As we will see from Theorem 2.4.2, $\frac{1}{8} + x^2$ is the minimax approximation in P_3 for $|x|$ on $[-1,1]$. Note that although $\frac{1}{8} + x^2$ is a polynomial of degree *two*, it is the minimax approximation to $|x|$ on $[-1,1]$ out of all polynomials in P_3. $\blacksquare$

Theorem 2.4.2 Let $f \in C[-1,1]$ and suppose there exists $p \in P_n$ such that $f - p$ equioscillates on $n+2$ points belonging to $[-1,1]$. Then p is the minimax approximation in P_n for f on $[-1,1]$.

Proof. Suppose p is not the minimax approximation. We will prove the theorem by showing that this assumption cannot be true. (We then say that we have obtained a *contradiction* to the initial assumption. This style of proof is called *reductio ad absurdum.*) If p is not the minimax approximation, there must exist some $q \in P_n$ such that $p + q \in P_n$ is the minimax approximation. Now let us compare the graphs of $f - p$, which equioscillates on $n+2$ points on $[-1,1]$, and $f - p - q$. Since the latter error curve must have the smaller maximum modulus, the effect of adding q to p must be to reduce the size of the modulus of the error function $f - p$ on all $n+2$

equioscillation points. In particular, this must mean that q alternates in sign on these $n+2$ points, and thus must have at least $n+1$ zeros. Since $q \in P_n$, this is impossible, contradicting our initial assumption that p is not the minimax approximation. ■

Example 2.4.2 For each $n \geq 0$, let us define

$$p(x) = x^{n+1} - \frac{1}{2^n} T_{n+1}(x),$$

so that $p \in P_n$. (See Problem 2.4.3 for a refinement of this observation.) Then, since

$$x^{n+1} - p(x) = \frac{1}{2^n} T_{n+1}(x)$$

equioscillates on $n+2$ points belonging to $[-1,1]$, it follows from Theorem 2.4.2 that p is the minimax approximation P_n for x^{n+1} on $[-1,1]$. From Theorem 2.2.5 we see that the same p is also the best approximation for x^{n+1} on $[-1,1]$ with respect to the weighted square norm, with weight $\omega(x) = (1-x^2)^{-1/2}$. Finally, if we write $\|\cdot\|_\infty$ to denote the maximum norm, defined by (2.5), we deduce from the minimax approximation for x^{n+1} that $\|(x-x_0)(x-x_1)\cdots(x-x_n)\|_\infty$ is minimized over all choices of the abscissas x_j by choosing the x_j as the zeros of T_{n+1}. Theorem 2.2.6 shows us that this also holds when we replace $\|\cdot\|_\infty$ by the weighted square norm, defined by (2.53) with weight function $(1-x^2)^{-1/2}$. ■

In Theorem 2.4.2 we showed that the equioscillation property is a sufficient condition for a minimax approximation. The next theorem, whose proof is a little harder, shows that the equioscillation property is also *necessary*, which is why we call it the characterizing property of minimax approximation.

Theorem 2.4.3 Let $f \in C[-1,1]$, and let $p \in P_n$ denote a minimax approximation for f on $[-1,1]$. Then there exist $n+2$ points on $[-1,1]$ on which $f-p$ equioscillates.

Proof. Let us write $E(x) = f(x) - p(x)$ and assume that E equioscillates on fewer than $n+2$ points. We will show that this assumption leads to a contradiction. We can exclude the case where E is the zero function, for then there is nothing to prove. Then we argue that E must equioscillate on at least two points, for otherwise, we could add a suitable constant to p so as to reduce the maximum modulus. Thus we can assume that E equioscillates on k points, where $2 \leq k < n+2$. Now let us choose $k-1$ points,

$$-1 < x_1 < x_2 < \cdots < x_{k-1} < 1,$$

so that there is one equioscillation point of E on each of the intervals

$$[-1, x_1), (x_1, x_2), \ldots, (x_{k-2}, x_{k-1}), (x_{k-1}, 1].$$

Now we construct a polynomial q such that

$$q(x) = \pm C(x - x_1)(x - x_2) \cdots (x - x_{k-1}),$$

with $C = 1/\|(x - x_1)(x - x_2) \cdots (x - x_{k-1})\|_\infty$. It follows from the choice of C that $\|q\|_\infty = 1$. Finally, we choose the plus or minus sign so that q takes the same sign as E on all points where $\|E\|_\infty$ is attained. Now let S_- denote the set of points on $[-1, 1]$ where $E(x)q(x) \leq 0$, and let S_+ denote the set of points on $[-1, 1]$ where $E(x)q(x) > 0$. Thus S_+ includes all points where $\|E\|_\infty$ is attained, and the union of S_- and S_+ is simply the interval $[-1, 1]$. We now define

$$d = \max_{x \in S_-} |E(x)| < \|E\|_\infty, \tag{2.133}$$

choose any $\theta > 0$, and define ξ as any point in $[-1, 1]$ for which

$$|E(\xi) - \theta q(\xi)| = \|E - \theta q\|_\infty. \tag{2.134}$$

Obviously, $\xi \in S_-$ or $\xi \in S_+$. If $\xi \in S_-$, it follows from (2.134) and (2.133) that

$$\|E - \theta q\|_\infty = |E(\xi)| + \theta |q(\xi)| \leq d + \theta. \tag{2.135}$$

On the other hand, if $\xi \in S_+$, it follows from (2.134) that

$$\|E - \theta q\|_\infty < \max\{|E(\xi)|, \theta |q(\xi)|\} \leq \max\{\|E\|_\infty, \theta\}. \tag{2.136}$$

In view of the definition of d in (2.133), let us now restrict the value of θ so that

$$0 < \theta < \|E\|_\infty - d.$$

Then, whether ξ is in S_- or in S_+, both (2.135) and (2.136) yield

$$\|E - \theta q\|_\infty < \|E\|_\infty.$$

This contradicts our assumption that $k < n + 2$, and completes the proof. Note that the polynomial q, defined above, is in P_{k-1}, and so we are able to force the above contradiction only if $k < n + 2$, so that $k - 1 \leq n$ and thus q is in P_n. ■

The uniqueness of a best approximation for the weighted square norm, discussed in the previous section, follows from the way we derived it as the unique solution of a system of linear equations. We now show how the uniqueness of a minimax approximation can be deduced by deploying the same ideas used above in the proof of Theorem 2.4.2.

Theorem 2.4.4 If $f \in C[-1, 1]$, there is a unique minimax approximation in P_n for f on $[-1, 1]$.

Proof. Let $p \in P_n$ denote a minimax approximation for f, and if this approximation is not unique, let $p + q \in P_n$ denote another minimax approximation. Then, using the same kind of argument as in the proof of Theorem 2.4.2, we argue that $q \in P_n$ must be alternately ≥ 0 and ≤ 0 on the $n + 2$ equioscillation points of $f - p$. We deduce that this is possible only when $q(x) \equiv 0$, which completes the proof. ■

We saw in the last section that the best approximation with respect to a weighted square norm is an interpolating polynomial for the given function. The above equioscillation theorems show that this is true also for minimax approximations, and we state this as a theorem.

Theorem 2.4.5 Let $f \in C[-1, 1]$ and let $p \in P_n$ denote the minimax polynomial for f. Then there exist $n+1$ points in $[-1, 1]$ on which p interpolates f. ■

It is convenient to introduce a new item of notation in the following definition.

Definition 2.4.2 We write

$$E_n(f) = \|f - p\|_\infty, \tag{2.137}$$

where $\|\cdot\|_\infty$ denotes the maximum norm on $[-1, 1]$, and $p \in P_n$ is the minimax approximation for $f \in C[-1, 1]$. ■

Theorem 2.4.5 leads to an estimate of the minimax error $E_n(f)$ that is similar to the estimate given in Theorem 2.2.7 for the error of a best approximation with respect to a weighted square norm.

Theorem 2.4.6 If $f \in C^{n+1}[-1, 1]$, then the error of the minimax polynomial $p \in P_n$ for f on $[-1, 1]$ satisfies

$$E_n(f) = \|f - p\|_\infty = \frac{1}{2^n} \frac{|f^{(n+1)}(\xi)|}{(n+1)!}, \tag{2.138}$$

where $\xi \in (-1, 1)$.

Proof. In the light of Theorem 2.4.5, let the minimax polynomial p interpolate f at the points $x_0, x_1, \ldots, x_n$ in $[-1, 1]$. Then, from (1.25),

$$f(x) - p(x) = (x - x_0)(x - x_1) \cdots (x - x_n) \frac{f^{(n+1)}(\xi_x)}{(n+1)!},$$

where $\xi_x \in (-1, 1)$, and thus

$$\|f - p\|_\infty \geq \|(x - x_0)(x - x_1) \cdots (x - x_n)\|_\infty \frac{\min |f^{(n+1)}(x)|}{(n+1)!}, \tag{2.139}$$

the minimum being over $[-1,1]$. It follows from our discussion in Example 2.4.2 that

$$\|(x-x_0)\cdots(x-x_n)\|_\infty \geq \|(x-x_0^*)\cdots(x-x_n^*)\|_\infty = \frac{1}{2^n},$$

where the x_j^* denote the zeros of T_{n+1}, and thus (2.139) yields

$$\|f-p\|_\infty \geq \frac{1}{2^n}\frac{\min|f^{(n+1)}(x)|}{(n+1)!}. \tag{2.140}$$

If p^* denotes the interpolating polynomial for f on the zeros of T_{n+1}, then, again using the error term for interpolation (1.25), we obtain the inequality

$$\|f-p^*\|_\infty \leq \|(x-x_0^*)(x-x_1^*)\cdots(x-x_n^*)\|_\infty \frac{\max|f^{(n+1)}(x)|}{(n+1)!},$$

and so

$$\|f-p\|_\infty \leq \|f-p^*\|_\infty \leq \frac{1}{2^n}\frac{\max|f^{(n+1)}(x)|}{(n+1)!}. \tag{2.141}$$

Finally, the theorem follows from (2.140), (2.141), and the continuity of $f^{(n+1)}$. ■

Example 2.4.3 If we apply Theorem 2.4.6 to the function e^x, we obtain

$$\frac{1}{2^n}\frac{e^{-1}}{(n+1)!} \leq \|e^x - p_n(x)\|_\infty \leq \frac{1}{2^n}\frac{e}{(n+1)!},$$

where p_n denotes the minimax polynomial in P_n for e^x on $[-1,1]$. For example, with $n=6$ and 7, we have the bounds

$$0.11\times 10^{-5} < \|e^x - p_6(x)\|_\infty < 0.85\times 10^{-5}$$

and

$$0.71\times 10^{-7} < \|e^x - p_7(x)\|_\infty < 0.53\times 10^{-6}. \quad ■$$

The next theorem shows that the minimax polynomial is not the only approximant to f that has an error term of the form (2.138).

Theorem 2.4.7 If p_n^* denotes the interpolating polynomial on the zeros of the Chebyshev polynomial T_{n+1} for $f \in C^{n+1}[-1,1]$, then

$$\|f-p_n^*\|_\infty = \frac{1}{2^n}\frac{|f^{(n+1)}(\eta)|}{(n+1)!}, \tag{2.142}$$

where $\eta \in (-1,1)$.

Proof. This is easily verified by adapting the proof of Theorem 2.4.6, and the details are left to the reader. ■

Example 2.4.4 Let p_n^* denote the interpolating polynomial for e^x on the zeros of the Chebyshev polynomial T_{n+1}. Then it follows from Theorem 2.4.7 that

$$\max_{-1\le x\le 1} |e^x - p_n^*(x)| = \frac{e^{\eta_n}}{2^n(n+1)!}, \tag{2.143}$$

where $-1 < \eta_n < 1$. It is clear that the sequence of polynomials (p_n^*) converges uniformly to e^x on $[-1,1]$. ■

We saw in Section 2.2 that when the weight function is even, the best weighted square norm approximation for a function f is even or odd, according as f is even or odd, respectively. We can deduce from the equioscillation property (see Problems 2.4.9 and 2.4.10) that a minimax approximation for a function f is even or odd, according as f is even or odd, respectively.

It is clear from Example 2.4.1 that $\frac{1}{8}+x^2$ is the minimax approximation in P_2 for $|x|$ on $[-1,1]$. However, we found that the error function equioscillates on five points, and so, by Theorem 2.4.2, $\frac{1}{8}+x^2$ is also the minimax approximation in P_3 for $|x|$, as stated in Example 2.4.1. This shows that if we seek a minimax polynomial $p_n \in P_n$ for a given function f, we could find that $f - p_n$ has more than $n+2$ equioscillation points. For example, the minimax polynomial in P_0 (a constant) for the function T_k on $[-1,1]$, with $k > 0$, is simply the zero function. In this case, the error function has $k+1$ equioscillation points. The following theorem gives a condition that ensures that $f - p_n$ has exactly $n+2$ equioscillation points.

Theorem 2.4.8 If $f \in C^{n+1}[-1,1]$ and $f^{(n+1)}$ has no zero in $(-1,1)$, then the error of the minimax polynomial $p \in P_n$ for f on $[-1,1]$ equioscillates on $n+2$ points, and on no greater number of points.

Proof. Suppose that $f - p$ has k equioscillation points in the interior of $[-1,1]$. Since there can be at most two equioscillation points at the endpoints $x = \pm 1$, Theorem 2.4.3 shows that $k \ge n$. This means that $f' - p'$ has at least $k \ge n$ zeros in $(-1,1)$. On applying Rolle's theorem, we deduce that $f'' - p''$ has at least $k-1$ zeros in $(-1,1)$. Since the nth derivative of p' is zero, the repeated application of Rolle's theorem n times to the function $f' - p'$ shows that $f^{(n+1)}$ has at least $k-n$ zeros in $(-1,1)$. Since, by our assumption, $f^{(n+1)}$ has no zero in $(-1,1)$, we deduce that $k = n$, and this completes the proof. ■

Apart from the approximation of x^{n+1} by a polynomial in P_n, and some simple cases involving low-order polynomial approximations for a few functions, we have said nothing about how to *compute* minimax approximations. We now discuss a class of algorithms based on the work of E. Ya. Remez (1896–1975) in the 1930s.

Algorithm 2.4.1 The following algorithm computes a sequence of polynomials that converges uniformly to the minimax approximation $p \in P_n$

for a given function $f \in C[-1,1]$. Corresponding to each polynomial in the sequence there is a set X of $n+2$ points. This set of points converges to a set of points in $[-1,1]$ on which the error $f-p$ equioscillates.

Step 1 Choose an initial set $X = \{x_1, x_2, \ldots, x_{n+2}\} \subset [-1,1]$.

Step 2 Solve the system of linear equations

$$f(x_j) - q(x_j) = (-1)^j e, \quad 1 \leq j \leq n+2,$$

to obtain a real number e and a polynomial $q \in P_n$.

Step 3 Change the set X (as described below) and go to Step 2 unless a "stopping criterion" has been met. ■

Let us denote a polynomial $q \in P_n$ that occurs in Step 2 of the above Remez algorithm by

$$q(x) = a_0 + a_1 x + \cdots + a_n x^n.$$

We now verify that, provided that the x_j are distinct, the system of linear equations in Step 2 has a nonsingular matrix, and so the linear system has a unique solution. The matrix associated with these equations is

$$\mathbf{A} = \begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^n & -e \\ 1 & x_2 & x_2^2 & \cdots & x_2^n & +e \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_{n+2} & x_{n+2}^2 & \cdots & x_{n+2}^n & (-1)^{n+2} e \end{bmatrix}.$$

Now, a necessary and sufficient condition for a matrix to be singular is that its columns be linearly dependent. If the columns of the above matrix are denoted by $\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_{n+2}$, and they are linearly dependent, then there exist real numbers $\lambda_1, \lambda_2, \ldots, \lambda_{n+2}$, not all zero, such that

$$\lambda_1 \mathbf{c}_1 + \lambda_2 \mathbf{c}_2 + \cdots + \lambda_{n+2} \mathbf{c}_{n+2} = \mathbf{0}, \tag{2.144}$$

where $\mathbf{0}$ is the zero column vector with $n+2$ elements. If we write

$$q(x) = \lambda_1 + \lambda_2 x + \cdots + \lambda_{n+1} x^n,$$

then the vector equation (2.144) is equivalent to the scalar system of linear equations

$$q(x_j) + \lambda_{n+2}(-1)^j e = 0, \quad 1 \leq j \leq n+2.$$

Since this would imply that $q \in P_n$ is alternately ≥ 0 and ≤ 0 on the $n+2$ distinct x_j, this is impossible, and we conclude that the above matrix $\mathbf{A}$ is nonsingular. Thus, provided that the x_j are distinct, the system of equations in Step 2 of the Remez algorithm always has a unique solution.

We usually choose the set X in Step 1 of the Remez algorithm to be the set of $n+2$ extreme points of T_{n+1}. Then we see from Example 2.4.2 that for the special case of $f(x) = x^{n+1}$, this choice of X immediately gives us the minimax polynomial when we solve the linear system in Step 2.

The solution of the linear system in Step 2 yields a polynomial $q \in P_n$ and some number e. It follows from Theorem 2.4.2 that if $\|f - q\|_\infty$ is sufficiently close to $|e|$, then q must be close to the minimax polynomial, and we would then terminate the algorithm. Otherwise, if ξ is a point such that $|f(\xi) - q(\xi)| = \|f - q\|_\infty$, we change the point set X by including ξ and deleting one of the existing points in X so that $f - q$ still alternates in sign on the new point set. In the scheme that we have just proposed, we change only one point each time we carry out Step 3. There is a second version of the algorithm, which converges more rapidly, in which we amend the point set X by including $n+2$ local extreme values of $f - q$ and delete existing points of X so that $f - q$ still alternates in sign on the new point set.

Suppose f is an even function and we seek its minimax approximation in P_{2n+1}. Let us choose as the initial set X the $2n+3$ extreme points of T_{2n+2}. Thus X consists of the abscissa $x = 0$ and $n+1$ pairs of abscissas of the form $\pm x_j$. The minimax polynomial p must be an even polynomial (see Problem 2.4.9), and is therefore of degree $2n$ or less. We can then see from the symmetry in the linear equations in Step 2 that the same is true of the polynomial q. The linear equations in Step 2 thus contain only $n+2$ coefficients, namely, e and the $n+1$ coefficients to determine p. Because of the symmetry in X, we need write down only the $n+2$ equations corresponding to the nonnegative abscissas in X. We can see that this simplification persists as we work through the algorithm. At each stage the set X is symmetric with respect to the origin, and the polynomial q is even. A corresponding simplification also occurs when we apply the Remez algorithm to an odd function.

Example 2.4.5 Let us use the Remez algorithm to compute the minimax polynomial in P_3 for $1/(1+x^2)$ on $[-1, 1]$. We require five equioscillation points, and initially in Step 1 we choose the set of extreme points of T_4,

$$X = \left\{-1, -\tfrac{1}{2}\sqrt{2}, 0, \tfrac{1}{2}\sqrt{2}, 1\right\}.$$

In Step 2 let us write $q(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3$. We see that the solution of the system of five linear equations is obtained by choosing $a_1 = a_3 = 0$ and solving the three equations corresponding to $x = 0$, $\frac{1}{2}\sqrt{2}$, and 1. This is in agreement with our remark above concerning the application of the Remez algorithm to an even function. We thus have the equations

$$\begin{aligned} 1 - a_0 &= e, \\ \frac{2}{3} - a_0 - \frac{1}{2}a_2 &= -e, \\ \frac{1}{2} - a_0 - a_2 &= e, \end{aligned}$$

whose solution is $a_0 = \frac{23}{24}$, $a_2 = -\frac{1}{2}$, and $e = \frac{1}{24} \approx 0.041667$. In this case, the error function has just one turning value in $(0, 1)$. Although we could now differentiate the error function to determine the turning value, this would not be in the spirit of the algorithm. Instead, we simply evaluate the error function on a suitably dense set of points on $[0, 1]$, say at intervals of 0.01. We find that the extreme values on this finite point set are as follows:

x	0	0.64	1
$f(x) - p(x)$	0.041667	−0.044112	0.041667

We now change the point set to $X = \{-1, -0.64, 0, 0.64, 1\}$ and go to Step 2. This time the solution of the linear system is $a_0 = 0.957111$, $a_2 = -\frac{1}{2}$, and $e = 0.042889$, to six decimal places. We evaluate the new error function on the same finite set of points and find that the extreme values are again at $x = 0, 0.64$, and 1, where $f - q$ assumes the values 0.042889, −0.042890, and 0.042889, respectively. Since the set X cannot be changed further, we accept the current polynomial q as being sufficiently close to the true minimax polynomial p, and terminate the algorithm. If we wish greater accuracy we can repeat the calculation, but evaluate the error at each stage on a finer grid of points. If we return to the original error function $f - q$ and evaluate it at intervals of 0.001 in the vicinity of 0.64, we find that this extreme point is more precisely given by 0.644, where the value of $f - q$ is −0.044120. Therefore, we now change the point set to $X = \{-1, -0.644, 0, 0.644, 1\}$ and go to Step 2. This time the solution of the linear system is $a_0 = 0.957107$, $a_2 = -\frac{1}{2}$, and $e = 0.042893$, to six decimal places. We evaluate the new error function at intervals of 0.01, and again refine the extreme point near 0.64. We find that the extreme values are

x	0	0.644	1
$f(x) - q(x)$	0.042893	−0.042893	0.042893

Thus we have the polynomial $0.957107 - 0.5x^2$ as a refined estimate of p, with minimax error 0.042893. In fact, for the function in this example, we can find p exactly (see Problem 2.4.7). This serves as a check on our calculations here. We can see that our last estimate of p above is correct to six decimal places. ■

We now state and prove a theorem, due to C. J. de la Vallée Poussin (1866–1962), which we can apply to give lower and upper bounds for the minimax error after each iteration of the Remez algorithm.

Theorem 2.4.9 Let $f \in C[-1,1]$ and $q \in P_n$. Then if $f - q$ alternates in sign on $n+2$ points

$$-1 \leq x_1 < x_2 < \cdots < x_{n+2} \leq 1,$$

we have

$$\min_j |f(x_j) - q(x_j)| \leq \|f - p\|_\infty \leq \|f - q\|_\infty, \tag{2.145}$$

where $p \in P_n$ denotes the minimax approximation for f on $[-1,1]$.

Proof. We need concern ourselves only with the left-hand inequality in (2.145), since the right-hand inequality follows immediately from the definition of p as the minimax approximation. Now let us write

$$(f(x_j) - q(x_j)) - (f(x_j) - p(x_j)) = p(x_j) - q(x_j). \tag{2.146}$$

If the left-hand inequality in (2.145) does *not* hold, then the right-hand inequality in the following line must hold:

$$|f(x_j) - p(x_j)| \leq \|f - p\|_\infty < |f(x_j) - q(x_j))|, \quad 1 \leq j \leq n+2.$$

The left hand inequality above is a consequence of the definition of the norm. It follows that the sign of the quantity on the left side of (2.146) is that of $f(x_j) - q(x_j)$, which alternates over the x_j. Thus $p - q \in P_n$ alternates in sign on $n+2$ points, which is impossible. This completes the proof. ■

The above theorem has the following obvious application to the Remez algorithm.

Theorem 2.4.10 At each stage in the Remez algorithm, we have

$$|e| \leq E_n(f) \leq \|f - q\|_\infty, \tag{2.147}$$

where $E_n(f)$ is the minimax error, and e and q are obtained from the solution of the linear equations in Step 2 of the algorithm.

Proof. From the way e and q are constructed in Step 2 of the algorithm, $f - q$ alternates in sign on the $n+2$ points x_j belonging to the set X, and

$$|e| = |f(x_j) - q(x_j)|, \quad 1 \leq j \leq n+2.$$

Hence (2.147) follows immediately from (2.145). ■

Let $e^{(i)}$ and $q^{(i)}$ denote the number e and the polynomial q that occur in the ith stage of the Remez algorithm. The sequence $\left(\left|e^{(i)}\right|\right)$ increases, and the sequence $\left(\left\|f - q^{(i)}\right\|_\infty\right)$ decreases, and in principle both sequences converge to the common limit $E_n(f)$. In practice, these limits are usually not attained, because we evaluate the error function at each stage at only a finite number of points. However, the inequalities (2.147) provide a reliable indicator of how close we are to the minimax polynomial at each stage.

Example 2.4.6 To illustrate the second version of the Remez algorithm, mentioned above, in which we amend the point set X by including $n+2$ local extreme values of $f-q$, let us find the minimax $p \in P_3$ for e^x on $[-1,1]$. We will take the initial set X as the set of extreme points of T_4, as we did in Example 2.4.5. The solution of the linear equations in Step 2 then yields, to six decimal places, $e = 0.005474$ and

$$q(x) = 0.994526 + 0.995682x + 0.543081x^2 + 0.179519x^3.$$

On evaluating $f-q$ at intervals of 0.01, we find that it has extreme values at $x = -1, -0.68, 0.05, 0.73$, and 1. The error $f-q$ has the value $e = 0.005474$ at $x = \pm 1$, by construction, and has the following values at the three interior extreme points:

x	-0.68	0.05	0.73
$f(x)-q(x)$	-0.005519	0.005581	-0.005537

If we now evaluate $f-q$ at intervals of 0.001 in the vicinity of the interior extreme points, we can refine these to give -0.683, 0.049, and 0.732. We therefore amend the set X, making three changes to give

$$X = \{-1, -0.683, 0.049, 0.732, 1\},$$

and repeat Step 2. This time we obtain $e = 0.005528$ and

$$q(x) = 0.994580 + 0.995668x + 0.542973x^2 + 0.179534x^3.$$

When we evaluate $f-p$ at intervals of 0.01, we find that it has extreme points at $x = -1$, -0.68, 0.05, 0.73, and 1, where it assumes the values ± 0.005528. We therefore accept this polynomial q as being sufficiently close to the minimax approximation p. The interior equioscillation points are more precisely -0.682, 0.050, and 0.732. ■

If a function is defined on a set $X = \{x_0, x_1, \ldots, x_N\}$, we can define a norm

$$\|f\| = \max_{0 \le j \le N} |f(x_j)|,$$

and seek a polynomial $p \in P_n$, with $n \le N$, that minimizes $\|f-p\|$. This is called a minimax approximation on the set X. Such minimax approximations behave much like those on a finite interval, as we have discussed in this section. For example, with $n < N$, the error function $f-p$ equioscillates on $n+2$ points in X. Also, when using a Remez algorithm, we can expect to locate a minimax polynomial on a finite interval only approximately, as we saw above, whereas we can find a minimax polynomial on a finite point set X precisely.

Problem 2.4.1 Show that the minimax polynomial $p \in P_1$ for $2/(x+3)$ on $[-1,1]$ is $p(x) = \frac{1}{2}\sqrt{2} - \frac{1}{4}x$, and find the minimax error.

Problem 2.4.2 Let $ax + b$ denote the minimax polynomial in P_1 for the function $1/x$ on the interval $[\alpha, \beta]$, where $0 < \alpha < \beta$. Show that

$$a = -\frac{1}{\alpha\beta} \quad \text{and} \quad b = \frac{1}{2}\left(\frac{1}{\sqrt{\alpha}} + \frac{1}{\sqrt{\beta}}\right)^2,$$

and that

$$\max_{\alpha \le x \le \beta} \left| ax + b - \frac{1}{x} \right| = \frac{1}{2}\left(\frac{1}{\sqrt{\alpha}} - \frac{1}{\sqrt{\beta}}\right)^2.$$

Verify that the minimax error is attained at α, β, and at the intermediate point $\sqrt{\alpha\beta}$.

Problem 2.4.3 Use the property that the Chebyshev polynomial T_n is an odd or even function when n is odd or even, respectively, to show that the polynomial p defined by

$$p(x) = x^{n+1} - \frac{1}{2^n}T_{n+1}(x)$$

is of degree $n-1$.

Problem 2.4.4 Let $p_n \in P_n$ denote the minimax polynomial for $\sin \frac{1}{2}\pi x$ on $[-1,1]$. Show that

$$0 < \| \sin \tfrac{1}{2}\pi x - p_n(x) \|_\infty \le \left(\frac{\pi}{4}\right)^{n+1} \frac{2}{(n+1)!}.$$

Problem 2.4.5 Let $f \in C[-1,1]$ and let

$$m = \min_{-1 \le x \le 1} f(x), \quad M = \max_{-1 \le x \le 1} f(x).$$

Show that $\frac{1}{2}(m+M) \in P_0$ is a minimax approximation for f on $[-1,1]$.

Problem 2.4.6 By first showing that the error function has turning values at $x = -\frac{2}{3}$ and $x = \frac{1}{3}$, show that the function $1/(3x+5)$ on $[-1,1]$ has the minimax approximation $\frac{1}{48}(6x^2 - 8x + 9)$ in P_2, and determine the minimax error.

Problem 2.4.7 Verify that the minimax approximation in P_3 for the function $1/(1+x^2)$ on $[-1,1]$ is $\frac{1}{4} + \frac{1}{2}\sqrt{2} - \frac{1}{2}x^2$, by showing that the error function equioscillates on the five points $0, \pm\xi$, and ± 1, where $\xi^2 = \sqrt{2} - 1$. Find the minimax error.

Problem 2.4.8 If $p \in P_n$ denotes the minimax polynomial for f on $[-1,1]$, show that $\lambda p + q$ is the minimax approximation for $\lambda f + q$, for any real λ and any $q \in P_n$. Thus show, using the result of Problem 2.4.6, that the minimax polynomial in P_2 for $(x+3)/(3x+5)$ is $(6x^2 - 8x + 21)/36$.

Problem 2.4.9 Let f be an even function on $[-1, 1]$, and let $p \in P_n$ denote the minimax approximation for f. By considering the equioscillation points, deduce that $p(-x)$ is the minimax approximation for $f(-x)$. Since $f(-x) = f(x)$ and the minimax approximation is unique, deduce that p is also an even function.

Problem 2.4.10 By adapting the argument used in the previous problem, show that a minimax approximation for an odd function is itself odd.

2.5 The Lebesgue Function

Recall equation (1.10),

$$p_n(x) = \sum_{i=0}^{n} f(x_i) L_i(x),$$

where p_n is the Lagrange form for the polynomial that interpolates f on the abscissas $x_0, x_1, \ldots, x_n$, and the fundamental polynomials L_i are defined by (1.9). If the values $f(x_i)$ all have errors of modulus not greater than $\epsilon > 0$, what can we say about the resulting size of the error in evaluating $p_n(x)$ at any point on the interval $[a, b]$? Suppose that instead of evaluating p_n, we evaluate p_n^*, where

$$p_n^*(x) = \sum_{i=0}^{n} f^*(x_i) L_i(x)$$

and

$$|f(x_i) - f^*(x_i)| \leq \epsilon, \quad 0 \leq i \leq n.$$

It follows that

$$|p_n(x) - p_n^*(x)| \leq \epsilon\, \lambda_n(x),$$

where

$$\lambda_n(x) = \sum_{i=0}^{n} |L_i(x)|, \quad a \leq x \leq b. \tag{2.148}$$

Hence we have

$$\max_{a \leq x \leq b} |p_n(x) - p_n^*(x)| \leq \epsilon\, \Lambda_n, \tag{2.149}$$

where

$$\Lambda_n = \max_{a \leq x \leq b} \lambda_n(x). \tag{2.150}$$

Thus errors in the function values $f(x_i)$ of modulus not greater than ϵ result in an error in the interpolating polynomial whose modulus is not greater than $\epsilon \Lambda_n$. We call Λ_n the *Lebesgue* constant and λ_n the Lebesgue function associated with the point set $\{x_0, x_1, \ldots, x_n\} \subset [a, b]$. These are named after Henri Lebesgue (1875–1941). Both the textbook by T. J. Rivlin [48] and the paper by Lev Brutman [6] contain valuable surveys on the Lebesgue functions.

Now let us consider an infinite triangular array of abscissas,

$$
\begin{array}{rlllll}
 & x_0^{(0)} & & & & \\
 & x_0^{(1)} & x_1^{(1)} & & & \\
X: & x_0^{(2)} & x_1^{(2)} & x_2^{(2)} & & \\
 & \vdots & & & & \\
 & x_0^{(n)} & x_1^{(n)} & \cdots & x_n^{(n)} & \\
 & \vdots & & & &
\end{array}
\tag{2.151}
$$

where for each $n \geq 0$,

$$
a \leq x_0^{(n)} < x_1^{(n)} < \cdots < x_n^{(n)} \leq b. \tag{2.152}
$$

For any function f defined on $[a, b]$, we can construct a sequence of polynomials $p_0,\ p_1,\ p_2, \ldots$, where p_n interpolates f on the set of abscissas $\left\{x_0^{(n)}, x_1^{(n)}, \ldots, x_n^{(n)}\right\}$, which forms the $(n+1)$th row of the above triangular array. When we wish to emphasize their dependence on the array X, we will write $\Lambda_n(X),\ n \geq 0$, to denote the Lebesgue constants on $[a, b]$ associated with this array of abscissas, and write the Lebesgue functions in the form $\lambda_n(X; x)$.

Example 2.5.1 In the above triangular array X let us choose $x_0^{(0)} = 0$, and for each $n \geq 1$, let

$$
x_i^{(n)} = -1 + \frac{2i}{n}, \qquad 0 \leq i \leq n.
$$

Thus the abscissas in each row of X are equally spaced on the interval $[-1, 1]$. Let us take $[a, b]$ in (2.152) to be $[-1, 1]$.

If we interpolate on the abscissas in the second row of X, namely $\{-1, 1\}$, the fundamental polynomials are $L_0(x) = \frac{1}{2}(1 - x)$ and $L_1(x) = \frac{1}{2}(1 + x)$. Thus the Lebesgue function on $[-1, 1]$ is

$$
\lambda_1(x) = \frac{1}{2}(1 - x) + \frac{1}{2}(1 + x) = 1, \quad -1 \leq x \leq 1,
$$

and the Lebesgue constant on $[-1, 1]$ is $\Lambda_1(X) = 1$.

Let us also obtain an explicit form for the Lebesgue function $\lambda_2(x)$ on $[-1, 1]$ for the set of interpolating abscissas in the third row of X, which is

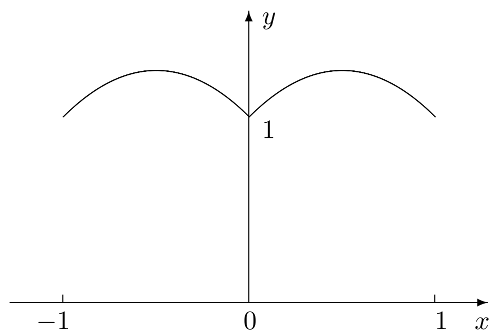

FIGURE 2.2. The Lebesgue function $\lambda_2(x)$ defined in Example 2.5.1.

$\{-1, 0, 1\}$. In this case the fundamental polynomials are

$$L_0(x) = \frac{1}{2}x(x-1), \quad L_1(x) = 1 - x^2, \quad L_1(x) = \frac{1}{2}x(x+1).$$

The Lebesgue function $\lambda_2(x)$ may be expressed as a quadratic polynomial in each of the intervals $[-1, 0]$ and $[0, 1]$. We find that

$$\lambda_2(x) = \frac{1}{2}x(x-1) + (1-x^2) - \frac{1}{2}x(1+x) = 1 - x - x^2, \quad -1 \leq x \leq 0,$$

and

$$\lambda_2(x) = -\frac{1}{2}x(x-1) + (1-x^2) + \frac{1}{2}x(1+x) = 1 + x - x^2, \quad 0 \leq x \leq 1.$$

It is easily verified that $\lambda_2(x) = 1$ for $x = 0, \pm 1$, and that $\lambda_2(x) \geq 1$ on $[-1, 1]$. We find that $\Lambda_2(X)$, the maximum value of $\lambda_2(x)$ on $[-1, 1]$ is $\frac{5}{4}$, and this value is attained at $x = \pm\frac{1}{2}$. See Figure 2.2. ■

Theorem 2.5.1 The Lebesgue function λ_n, defined by (2.148), is continuous on $[a, b]$ and is a polynomial of degree at most n on each of the subintervals $[a, x_0]$, $[x_0, x_1]$, $[x_1, x_2], \ldots, [x_{n-1}, x_n]$, and $[x_n, b]$. We also have

$$\lambda_n(x_i) = 1, \quad 0 \leq i \leq n, \tag{2.153}$$

and

$$\lambda_n(x) \geq 1, \quad a \leq x \leq b. \tag{2.154}$$

Proof. Equation (2.153) follows immediately from (2.148). To verify the inequality (2.154) we deduce from (1.42) that

$$1 = |L_0(x) + L_1(x) + \cdots + L_n(x)| \leq |L_0(x)| + |L_1(x)| + \cdots + |L_n(x)| = \lambda_n(x)$$

for $a \leq x \leq b$. Each function $|L_i(x)|$, and therefore the Lebesgue function λ_n, is obviously a polynomial of degree at most n on each of the intervals

$[a, x_0]$, $[x_0, x_1]$, $[x_1, x_2], \ldots, [x_{n-1}, x_n]$, and $[x_n, b]$. Also, each $|L_j(x)|$ is continuous on $[a, b]$, and thus the Lebesgue function λ_n is continuous on $[a, b]$. This completes the proof. ■

We now come to the main result of this section.

Theorem 2.5.2 Let $p_n \in P_n$ denote the minimax approximation for a given function $f \in C[a, b]$, and let $p_n^* \in P_n$ denote the interpolating polynomial for f on the abscissas $x_i^{(n)}$ in the $(n+1)$th row of the triangular array X defined in (2.151). Then

$$\|f - p_n^*\|_\infty \leq (1 + \Lambda_n(X))E_n(f), \tag{2.155}$$

where $\|\cdot\|_\infty$ denotes the maximum norm on $[a, b]$ and $E_n(f)$ is the minimax error, defined in (2.137).

Proof. We have from (1.10) that

$$p_n^*(x) = \sum_{i=0}^{n} f(x_i) L_i^{(n)}(x),$$

where $L_i^{(n)}$ is the fundamental polynomial that takes the value 1 at the abscissa $x_i^{(n)}$ and is zero on all abscissas $x_j^{(n)}$ with $j \neq i$. It then follows from the uniqueness of the interpolating polynomial that

$$p_n(x) = \sum_{i=0}^{n} p_n(x_i) L_i^{(n)}(x),$$

since this relation holds for any polynomial in P_n. On subtracting the last two equations, we immediately obtain

$$p_n^*(x) - p_n(x) = \sum_{i=0}^{n} (f(x_i) - p_n(x_i)) L_i^{(n)}(x),$$

and thus

$$|p_n^*(x) - p_n(x)| \leq \lambda_n(x) \max_{0 \leq i \leq n} |f(x_i) - p_n(x_i)|,$$

where λ_n is the Lebesgue function. We deduce that

$$\|p_n - p_n^*\| = \|p_n^* - p_n\| \leq \Lambda_n(X)E_n(f). \tag{2.156}$$

Let us now write

$$f(x) - p_n^*(x) = (f(x) - p_n(x)) + (p_n(x) - p_n^*(x)),$$

from which we can derive the inequality

$$\|f - p_n^*\| \leq \|f - p_n\| + \|p_n - p_n^*\|,$$

and we immediately have

$$\|f - p_n^*\| \leq (1 + \Lambda_n(X))E_n(f),$$

which completes the proof. ■

Theorem 2.5.3 The Lebesgue constants

$$\Lambda_n(X) = \max_{a \leq x \leq b} \sum_{i=0}^{n} \left| L_i^{(n)}(x) \right|, \quad n \geq 1,$$

are unchanged if we carry out a linear transformation $x = \alpha t + \beta$, with $\alpha \neq 0$, and the triangular array of interpolating abscissas $X = \left(x_i^{(n)} \right)$ is mapped to the triangular array $T = \left(t_i^{(n)} \right)$, where $x_i^{(n)} = \alpha t_i^{(n)} + \beta$.

Proof. Let the interval $a \leq x \leq b$ be mapped to $c \leq t \leq d$ under the linear transformation $x = \alpha t + \beta$, and let the fundamental polynomial $L_i^{(n)}(x)$ be mapped to

$$M_i^{(n)}(t) = \prod_{j \neq i} \left(\frac{t - t_j^{(n)}}{t_i^{(n)} - t_j^{(n)}} \right),$$

so that the Lebesgue function $\lambda_n(x)$ is mapped to

$$\lambda_n^*(t) = \sum_{i=0}^{n} \left| M_i^{(n)}(t) \right|.$$

Finally, we define

$$\Lambda_n^*(T) = \max_{c \leq t \leq d} \lambda_n^*(t).$$

We can verify that

$$\frac{x - x_j^{(n)}}{x_i^{(n)} - x_j^{(n)}} = \frac{t - t_j^{(n)}}{t_i^{(n)} - t_j^{(n)}},$$

and hence $M_i^n(t) = L_i^{(n)}(x)$. Consequently,

$$\Lambda_n^*(T) = \max_{c \leq t \leq d} \lambda_n^*(t) = \max_{a \leq t \leq b} \lambda_n(x) = \Lambda_n(X),$$

and this completes the proof. ■

The following result shows that the Lebesgue constants are not increased if we include the endpoints a and b in the set of interpolating abscissas.

Theorem 2.5.4 Consider an infinite triangular array X, as defined in (2.151), where

$$a \leq x_0^{(n)} < x_1^{(n)} < \cdots < x_n^{(n)} \leq b, \quad n \geq 0.$$

Now define an infinite triangular array T, where $t_0^{(0)}$ is defined arbitrarily, and for each $n > 0$, the elements in the $(n+1)$th row of T satisfy

$$x_i^{(n)} = \alpha_n t_i^{(n)} + \beta_n, \quad 0 \le i \le n.$$

In keeping with this transformation, we define

$$M_i^{(n)}(t) = L_i^{(n)}(x), \quad \text{where} \quad x = \alpha_n t + \beta_n, \quad 0 \le i \le n.$$

Then, if α_n and β_n are chosen so that $t_0^{(n)} = a$ and $t_n^{(n)} = b$, we have

$$\Lambda_n(T) = \max_{a \le t \le b} |\lambda_n^*(t)| = \max_{x_0^{(n)} \le x \le x_n^{(n)}} |\lambda_n(x)| \le \max_{a \le x \le b} |\lambda_n(x)| = \Lambda_n(X).$$

Proof. The proof of this theorem is on the same lines as that of Theorem 2.5.3. Note that since

$$x_0^{(n)} = \alpha_n t_0^{(n)} + \beta_n \quad \text{and} \quad x_n^{(n)} = \alpha_n t_n^{(n)} + \beta_n,$$

we need to choose

$$\alpha_n = \frac{x_n^{(n)} - x_0^{(n)}}{b-a} \quad \text{and} \quad \beta_n = \frac{b x_0^{(n)} - a x_n^{(n)}}{b-a}. \qquad \blacksquare$$

Example 2.5.2 Let us consider the Lebesgue constants $\Lambda_n(X)$ for interpolation at equally spaced abscissas. For convenience, we will work on the interval $[0,1]$, rather than on $[-1,1]$, as we did in Example 2.5.1. Thus the abscissas in the triangular array X are $x_0^{(0)} = \frac{1}{2}$ and

$$x_i^{(n)} = \frac{i}{n}, \qquad 0 \le i \le n.$$

Let us evaluate the Lebesgue function $\lambda_n(x)$ at

$$x = \xi_n = \frac{2n-1}{2n},$$

the midpoint of the subinterval $[\frac{n-1}{n}, 1]$. We already know that $\lambda_n(x)$ has the value 1 at the endpoints of this subinterval. A little calculation shows that

$$|L_i(\xi_n)| = \frac{1}{|2n-2i-1|} \cdot \frac{(2n-1)(2n-3)\cdots 3 \cdot 1}{2^n i!(n-i)!}.$$

We can simplify this last equation to give

$$|L_i(\xi_n)| = \frac{1}{|2n-2i-1|} \cdot \frac{1}{2^{2n}} \binom{2n}{n} \binom{n}{i}, \tag{2.157}$$

from which we derive the inequality

$$|L_i(\xi_n)| \geq \frac{1}{2n-1} \cdot \frac{1}{2^{2n}} \left(\begin{array}{c} 2n \\ n \end{array} \right) \left(\begin{array}{c} n \\ i \end{array} \right). \tag{2.158}$$

If we now sum the above inequality over i, and evaluate the binomial expansion of $(1+x)^n$ at $x = 1$ to give

$$\sum_{i=0}^{n} \left(\begin{array}{c} n \\ i \end{array} \right) = 2^n,$$

we see from (2.148) that

$$\lambda_n(\xi_n) > \frac{1}{2n-1} \cdot \frac{1}{2^n} \left(\begin{array}{c} 2n \\ n \end{array} \right), \quad \text{where} \quad \xi_n = \frac{2n-1}{2n}. \tag{2.159}$$

Note that the above inequality holds strictly because we have equality in (2.158) only for $i = 0$. Thus

$$\lambda_n(\xi_n) > \frac{\mu_n}{2n-1}, \quad \text{where} \quad \xi_n = \frac{2n-1}{2n},$$

and μ_n is defined in (2.29). If we now apply Stirling's formula, as we did to estimate μ_n in Problem 2.1.12, we find that

$$\Lambda_n(X) \geq \frac{1}{2n-1} \cdot \frac{1}{2^n} \left(\begin{array}{c} 2n \\ n \end{array} \right) \sim \frac{2^{n-1}}{\sqrt{\pi}\, n^{3/2}}. \tag{2.160}$$

In view of the factor 2^{n-1} on the right of (2.160), we have proved that the Lebesgue constants for equally spaced abscissas tend to infinity at least exponentially with n. ■

It is natural to seek an infinite triangular array of interpolating abscissas X that gives the smallest values of $\Lambda_n(X)$ for every value of n. It would be nice if for such an *optimal* array X, the sequence $(\Lambda_n(X))_{n=0}^{\infty}$ were bounded. However, this is not so, as the following result of Paul Erdős (1913–1996) shows.

Theorem 2.5.5 There exists a positive constant c such that

$$\Lambda_n(X) > \frac{2}{\pi} \log n - c, \tag{2.161}$$

for *all* infinite triangular arrays X. ■

For a proof, see [18]. A simple proof of a slightly weaker version of this theorem, that

$$\Lambda_n(X) > \frac{2}{\pi^2} \log n - 1$$

for every triangular array X, is given in Rivlin [48].

Thus, for every choice of the array X, the sequence $(\Lambda_n(X))_{n=0}^{\infty}$ grows at least as fast as $\log n$. Moreover, as we will see below, interpolation at the zeros of the Chebyshev polynomials yields a sequence of Lebesgue constants that grows only logarithmically with n and so is close to the optimal choice of abscissas. It is therefore clear that as measured by the rate of growth of the Lebesgue constants, interpolation at the zeros of the Chebyshev polynomials is substantially superior to interpolation at equally spaced abscissas.

Let T denote the infinite triangular array, as depicted in (2.151), whose $(n+1)$th row consists of the zeros of T_{n+1}, which we will write as

$$x_i^{(n)} = \cos\theta_i, \quad \text{where} \quad \theta_i = \frac{(2n+1-2i)\pi}{2n+2}, \quad 0 \le i \le n, \tag{2.162}$$

so that

$$-1 < x_0^{(n)} < x_1^{(n)} < \cdots < x_n^{(n)} < 1.$$

We can deduce from the result in Problem 1.1.3 that the fundamental polynomial $L_i^{(n)}$ can be expressed in the form

$$L_i^{(n)}(x) = \frac{T_{n+1}(x)}{\left(x - x_i^{(n)}\right) T'_{n+1}\left(x_i^{(n)}\right)}. \tag{2.163}$$

From the expression for the derivative of a Chebyshev polynomial given in Problem 2.2.11, we can see that

$$T'_{n+1}(x_i^{(n)}) = \frac{(n+1)\sin(n+1)\theta_i}{\sin\theta_i} = (-1)^{n-i}\left(\frac{n+1}{\sin\theta_i}\right),$$

and hence, with $x = \cos\theta$,

$$L_i^{(n)}(x) = (-1)^{n-i}\frac{\cos(n+1)\theta}{n+1} \cdot \frac{\sin\theta_i}{\cos\theta - \cos\theta_i}, \tag{2.164}$$

so that

$$\lambda_n(T;x) = \frac{|\cos(n+1)\theta|}{n+1}\sum_{i=0}^{n}\frac{\sin\theta_i}{|\cos\theta - \cos\theta_i|}. \tag{2.165}$$

S. N. Bernstein (1880–1968) obtained the asymptotic estimate (see [4])

$$\Lambda_n(T) \sim \frac{2}{\pi}\log(n+1), \tag{2.166}$$

and D. L. Berman [2] obtained the upper bound

$$\Lambda_n(T) < 4\sqrt{2} + \frac{2}{\pi}\log(n+1), \tag{2.167}$$

which, together with Bernstein's asymptotic estimate, tells us more about $\Lambda_n(T)$. Luttmann and Rivlin [36] conjectured that $\Lambda_n(T) = \lambda_n(T;1)$, and showed that

$$\lim_{n\to\infty}\left[\lambda_n(T;1) - \frac{2}{\pi}\log(n+1)\right] = \frac{2}{\pi}\left(\gamma + \log\frac{8}{\pi}\right) \approx 0.9625, \tag{2.168}$$

where

$$\gamma = \lim_{n\to\infty}\left[1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} - \log n\right] \approx 0.5772$$

is Euler's constant. The theorem that follows, which verifies Luttmann and Rivlin's conjecture, is due to Ehlich and Zeller [16] (see also Powell [46]). The proof given here is based on that given by Rivlin [48].

Theorem 2.5.6 The Lebesgue constants for the triangular array T, whose $(n+1)$th row consists of the zeros of T_{n+1}, are given by

$$\Lambda_n(T) = \lambda_n(T;1) = \frac{1}{n+1}\sum_{i=0}^{n}\cot\frac{\theta_i}{2} = \frac{1}{n+1}\sum_{i=0}^{n}\tan\frac{(2i+1)\pi}{4(n+1)}, \tag{2.169}$$

where θ_i is defined in (2.162).

Proof. We begin by rewriting (2.164) in the form (see Problem 2.5.2)

$$L_i^{(n)}(x) = (-1)^{n-i}\frac{\cos(n+1)\theta}{2(n+1)}\left(\cot\frac{\theta+\theta_i}{2} - \cot\frac{\theta-\theta_i}{2}\right),$$

and thus obtain

$$\lambda_n(T;x) = \frac{|\cos(n+1)\theta|}{2(n+1)}\sum_{i=0}^{n}\left|\cot\frac{\theta+\theta_i}{2} - \cot\frac{\theta-\theta_i}{2}\right|, \tag{2.170}$$

where $x = \cos\theta$. Let us write

$$C(\theta) = |\cos(n+1)\theta|\sum_{i=0}^{n}\left|\cot\frac{\theta+\theta_i}{2} - \cot\frac{\theta-\theta_i}{2}\right|. \tag{2.171}$$

Since $C(\theta) = C(-\theta)$, this is an even function that is periodic, with period 2π, and is positive for $-\infty < \theta < \infty$. If we replace θ by $\theta - k\pi/(n+1)$, where $1 \le k \le n+1$, then (see Problem 2.5.3) the only change that occurs on the right side of (2.171) is that the $2n+2$ cotangents are paired differently.

Let the maximum value of $C(\theta)$ on $[0,\pi]$ be attained at $\theta = \theta'$ such that

$$\frac{(2m-1)\pi}{2n+2} < \theta' \le \frac{(2m+1)\pi}{2n+2}, \tag{2.172}$$

where $0 \le m \le n$. We see from (2.172) that θ' belongs to an interval of width $\pi/(n+1)$ with midpoint $m\pi/(n+1)$, and to a smaller interval when $m = 0$ or n. In either case, we have

$$\left|\theta' - \frac{m\pi}{n+1}\right| \le \frac{\pi}{2(n+1)}.$$

Let us now define $\theta'' = |\theta' - m\pi/(n+1)|$, and then we have

$$0 \le \theta'' \le \frac{\pi}{2(n+1)}. \tag{2.173}$$

We know that $C(\theta'') = C(-\theta'')$ and that the expressions for $C(\theta'')$ and $C(\theta')$ differ only in the way the $2n+2$ cotangents are paired in (2.171).

It follows from (2.162) and (2.173) that

$$0 \le \frac{\theta'' + \theta_i}{2} \le \frac{\pi}{2} \quad \text{and} \quad 0 \le -\frac{\theta'' - \theta_i}{2} \le \frac{\pi}{2}, \quad 0 \le i \le n,$$

and thus

$$\cot \frac{\theta'' + \theta_i}{2} \ge 0 \quad \text{and} \quad \cot \frac{\theta'' - \theta_i}{2} \le 0, \quad 0 \le i \le n.$$

The latter inequalities make it clear that if we put $\theta = \theta''$ in (2.171), the resulting expression will have the same value for *any* permutation of the $2n+2$ cotangents. This shows that

$$\max_{0 \le \theta \le \pi} C(\theta) = C(\theta') = C(\theta'') = \max_{0 \le \theta \le \pi/(2n+2)} C(\theta).$$

To pursue this further, we see from (2.165), (2.170), and (2.171) that

$$C(\theta) = 2|\cos(n+1)\theta| \sum_{i=0}^{n} \frac{\sin\theta_i}{|\cos\theta - \cos\theta_i|}.$$

Then, for $0 \le \theta \le \pi/(2n+2)$, we have

$$C(\theta) = 2\sum_{i=0}^{n} \sin\theta_i \frac{\cos(n+1)\theta}{\cos\theta - \cos\theta_i},$$

and we see from Problem 2.5.4 that

$$C(\theta) = 2^{n+1} \sum_{i=0}^{n} \sin\theta_i \prod_{j \ne i} (\cos\theta - \cos\theta_j). \tag{2.174}$$

Since each factor $\cos\theta - \cos\theta_j$ is a nonnegative function of θ that decreases monotonically on $[0, \pi/(2n+2)]$, we see immediately that the maximum value of $C(\theta)$ is attained at $\theta = 0$. On comparing (2.171) and (2.170), and putting $\theta = 0$, we see that

$$\Lambda_n(T) = \lambda_n(T;1) = \frac{1}{n+1} \sum_{i=0}^{n} \cot\frac{\theta_i}{2} = \frac{1}{n+1} \sum_{i=0}^{n} \cot\frac{(2n+1-2i)\pi}{4(n+1)},$$

and (2.169) follows. This completes the proof. ■

It is not hard to deduce from (2.169) that the Lebesgue constants $\Lambda_n(T)$ grow logarithmically with n. Let

$$g(\theta) = \tan\theta - \theta,$$

so that $g(0) = 0$, and since

$$g'(\theta) = \sec^2\theta - 1 = \tan^2\theta,$$

we see that $g'(\theta) > 0$ for $0 < \theta < \frac{\pi}{2}$. It follows that

$$0 < \theta < \tan\theta, \qquad 0 < \theta < \frac{\pi}{2},$$

and thus

$$\frac{1}{\theta} - \cot\theta = \frac{1}{\theta} - \frac{1}{\tan\theta} > 0, \qquad 0 < \theta < \frac{\pi}{2}. \tag{2.175}$$

Let us now return to (2.169) and write

$$\Lambda_n(T) = \frac{1}{n+1}\sum_{i=0}^{n} \cot\frac{\theta_i}{2} = \Sigma_n - \Sigma_n^*, \tag{2.176}$$

say, where

$$\Sigma_n = \frac{1}{n+1}\sum_{i=0}^{n}\frac{2}{\theta_i} \qquad \text{and} \qquad \Sigma_n^* = \frac{1}{n+1}\sum_{i=0}^{n}\left(\frac{2}{\theta_i} - \cot\frac{\theta_i}{2}\right). \tag{2.177}$$

The inequality in (2.175) implies that $\Sigma_n^* > 0$, and it follows from (2.176) that

$$\Lambda_n(T) < \Sigma_n = \frac{1}{n+1}\sum_{i=0}^{n}\frac{2}{\psi_i}, \tag{2.178}$$

where

$$\psi_i = \theta_{n-i} = \frac{(2i+1)\pi}{2n+2},$$

giving the simple inequality

$$\Lambda_n(T) < \frac{4}{\pi}\sum_{i=0}^{n}\frac{1}{2i+1}.$$

On writing

$$S_n = \sum_{i=1}^{n}\frac{1}{i},$$

and using the result in Problem 2.5.5, we see that

$$\Lambda_n(T) < \frac{4}{\pi}(S_{2n+2} - \frac{1}{2}S_{n+1}) < \frac{4}{\pi}\left(1 + \log(2n+2) - \frac{1}{2}\log(n+2)\right).$$

Then, on applying the inequality in Problem 2.5.6, we can deduce that

$$\Lambda_n(T) < \frac{2}{\pi}\log n + \frac{4}{\pi}\left(1 + \frac{1}{2}\log\frac{16}{3}\right),$$

and thus

$$\Lambda_n(T) < \frac{2}{\pi}\log n + 3. \tag{2.179}$$

Having obtained this inequality for the Lebesgue constant $\Lambda_n(T)$, let us estimate how much we "gave away" when we discarded the positive quantity Σ_n^* from (2.176) to obtain the inequality in (2.178). An inspection of the expression for Σ_n^* in (2.177) reveals that

$$\lim_{n\to\infty} \Sigma_n^* = \frac{2}{\pi}\int_0^{\pi/2}\left(\frac{1}{\theta} - \cot\theta\right)d\theta = \frac{2}{\pi}\log(\pi/2) \approx 0.287. \tag{2.180}$$

In fact, Σ_n^* is the quantity obtained by applying the midpoint rule in composite form with $n+1$ subintervals to estimate the integral given as the limit of the sequence (Σ_n^*) in (2.180). (See Section 3.1.) A more detailed analysis can confirm what is suggested by the foregoing material, that Σ_n^* is smaller than the error incurred in estimating Σ_n.

Let T' denote the infinite triangular array whose $(n+1)$th row consists of the extreme points of the Chebyshev polynomial T_n, and thus

$$x_i^{(n)} = \cos\phi_{n-i}, \quad \text{where} \quad \phi_i = \cos\frac{i}{n}, \quad 0 \le i \le n. \tag{2.181}$$

Intuitively, in view of Theorem 2.5.4, one would expect the Lebesgue constants $\Lambda_n(T')$ to be small, since the abscissas in T' are distributed in a similar fashion to those in T, *and* each row of the array T' contains the endpoints ± 1. Using the methods employed in the proof of Theorem 2.5.6, it is not hard to show that

$$\lambda_n(T'; x) = \frac{|\sin n\theta|}{2n}\sum_{i=0}^{n}{}'' \left|\cot\frac{\theta+\phi_i}{2} + \cot\frac{\theta-\phi_i}{2}\right|, \tag{2.182}$$

where $x = \cos\theta$ and $\sum''$ denotes a sum whose first and last terms are halved. It is also not difficult to verify that $\Lambda_n(T') \le \Lambda_{n-1}(T)$ for all $n \ge 2$, with equality when n is odd. More precisely, we have

$$\Lambda_n(T') = \Lambda_{n-1}(T) = \frac{1}{n}\sum_{i=1}^{n}\tan\frac{(2i-1)\pi}{4n}, \quad n \text{ odd}, \tag{2.183}$$

and

$$\sum_{i=2}^{n}\tan\frac{(2i-1)\pi}{4n} < \Lambda_n(T') < \frac{1}{n}\sum_{i=1}^{n}\tan\frac{(2i-1)\pi}{4n}, \quad n \text{ even}. \tag{2.184}$$

n	$\Lambda_n(T)$	$\Lambda_n(T')$	$\Lambda_n(T^*)$
1	1.414	1.000	1.000
2	1.667	1.250	1.250
3	1.848	1.667	1.430
4	1.989	1.799	1.570
5	2.104	1.989	1.685
6	2.202	2.083	1.783
7	2.287	2.202	1.867
8	2.362	2.275	1.942
9	2.429	2.362	2.008
10	2.489	2.421	2.069
20	2.901	2.868	2.479
50	3.466	3.453	3.043
100	3.901	3.894	3.478
200	4.339	4.336	3.916
500	4.920	4.919	4.497
1000	5.361	5.360	4.937

TABLE 2.1. Comparison of the Lebesgue constants $\Lambda_n(T)$, $\Lambda_n(T')$, and $\Lambda_n(T^*)$.

Let T^* denote the infinite triangular array that is derived from T, the array whose rows consist of the zeros of the Chebyshev polynomials, by dividing the numbers in the $(n+1)$th row of T by $\cos(\pi/(2n+2)$, for $n \geq 1$. To be definite, we will choose the sole number in the first row of T^* as zero. We call the numbers in T^*, from the second row onwards, the *stretched* zeros of the Chebyshev polynomials. Thus, if the numbers in the $(n+1)$th row of T^* are denoted by $\bar{x}_i^{(n)}$, we see that

$$\bar{x}_0^{(n)} = -1 \quad \text{and} \quad \bar{x}_0^{(n)} = 1.$$

We know from Theorem 2.5.4 that we must have $\Lambda_n(T^*) \leq \Lambda_n(T)$, for $n \geq 1$, and Table 2.1 strongly suggests that

$$\Lambda_n(T) < \Lambda_n(T') < \Lambda_n(T^*)$$

for $n > 2$.

Luttmann and Rivlin [36] show that for every triangular array X whose abscissas satisfy

$$-1 = x_0^{(n)} < x_1^{(n)} < \cdots < x_n^{(n)} = 1, \quad n \geq 1,$$

each Lebesgue function $\lambda_n(X;x)$ has a single maximum

$$M_{j,n}(X) = \max_{x_{j-1}^{(n)} \leq x \leq x_j^{(n)}} \lambda_n(X;x), \quad 1 \leq j \leq n,$$

on each of the n subintervals $\left[x_{j-1}^{(n)}, x_j^{(n)}\right]$. As early as 1931 Bernstein [4] conjectured that if there existed an array X for which

$$M_{1,n}(X) = M_{2,n}(X) = \cdots = M_{n,n}(X), \quad n \geq 2, \tag{2.185}$$

then the array X would be optimal. Subsequently, in 1958, Paul Erdős [17] further conjectured that there is a unique array X^* for which (2.185) holds, and that for *any* array X,

$$\min_{1 \leq j \leq n} M_{j,n}(X) \leq \Lambda_n(X^*). \tag{2.186}$$

In 1978 two papers were published consecutively in the same issue of the *Journal of Approximation Theory*, one by T. A. Kilgore [29] and the other by C. de Boor and A. Pinkus [12], in which these conjectures were proved. It was shown by Lev Brutman [6] that there is little variation in the n numbers $M_{j,n}(T^*)$, since

$$\Lambda_n(T^*) \leq \min_{1 \leq j \leq n} M_{j,n}(T^*) + 0.201,$$

and thus

$$\Lambda_n(T^*) \leq \Lambda_n(X^*) + 0.201.$$

We may conclude that even if we do not know the optimal array X^* explicitly, it suffices for all practical purposes to use the array of stretched Chebyshev zeros T^*.

Example 2.5.3 Let us find the Lebesgue function $\lambda_3(x)$ on $[-1, 1]$ based on the set of interpolating abscissas $\{-1, -t, t, 1\}$, where $0 < t < 1$. With a little simplification, we find that

$$\lambda_3(x) = \frac{|x^2 - t^2|}{1 - t^2} + \frac{(1 - x^2)|t - x|}{2t(1 - t^2)} + \frac{(1 - x^2)|t + x|}{2t(1 - t^2)},$$

so that

$$\lambda_3(x) = \begin{cases} \dfrac{1 + t^2 - 2x^2}{1 - t^2}, & 0 \leq x \leq t, \\[2ex] \dfrac{-t^3 + x + tx^2 - x^3}{t(1 - t^2)}, & t < x \leq 1, \end{cases}$$

and $\lambda_3(x)$ is *even* on $[-1, 1]$. It is obvious that $\lambda_3(x)$ has a local maximum value at $x = 0$, say $M(t)$, and that

$$M(t) = \frac{1 + t^2}{1 - t^2}.$$

With a little more effort, we find that $\lambda_3(x)$ has local maximum values, say $M^*(t)$, at $x = \pm x^*(t)$, where

$$x^*(t) = \frac{1}{3}\left[t + (t^2 + 3)^{1/2}\right] \quad \text{for} \quad 0 < t < 1.$$

We can verify that $t < x^*(t) < 1$, and that

$$M^*(t) = \frac{9t - 25t^3 + (t^2+3)^{1/2}(6+2t^2)}{27t(1-t^2)}.$$

On the interval $0 < t < 1$, $M(t)$ increases monotonically and $M^*(t)$ decreases monotonically, and we find that

$$M(t) = M^*(t) \approx 1.423 \quad \text{for} \quad t \approx 0.4178.$$

Thus, as we see from the line above and from Table 2.1,

$$\Lambda_3(X^*) \approx 1.423 < \Lambda_3(T^*) \approx 1.430,$$

where X^* and T^* denote respectively the optimal array and the array of stretched Chebyshev zeros. ■

Problem 2.5.1 Show that the Lebesgue function $\lambda_1(x)$ on the interval $[-1, 1]$ for interpolation on the abscissas $\pm t$, where $0 < t \leq 1$, is given by

$$\lambda_1(x) = \begin{cases} -x/t, & -1 \leq x \leq -t, \\ 1, & -t < x \leq t, \\ x/t, & t < x \leq 1, \end{cases}$$

and so verify the value given for $\Lambda_1(T)$ in Table 2.1.

Problem 2.5.2 Write

$$\cot\frac{\theta+\theta_i}{2} - \cot\frac{\theta-\theta_i}{2} = \frac{\sin\frac{\theta-\theta_i}{2}\cos\frac{\theta+\theta_i}{2} - \sin\frac{\theta+\theta_i}{2}\cos\frac{\theta-\theta_i}{2}}{\sin\frac{\theta+\theta_i}{2}\sin\frac{\theta-\theta_i}{2}},$$

and hence show that

$$\cot\frac{\theta+\theta_i}{2} - \cot\frac{\theta-\theta_i}{2} = \frac{2\sin\theta_i}{\cos\theta - \cos\theta_i}.$$

Problem 2.5.3 Following Rivlin [48], let us write

$$c_i^+ = \cot\frac{\theta+\phi_i}{2} \quad \text{and} \quad c_i^- = \cot\frac{\theta-\phi_i}{2},$$

where $\phi_i = (2i+1)\pi/(2n+2)$, for $0 \leq i \leq n$. Show that if θ is replaced by $\theta - k\pi/(n+1)$, where $1 \leq k \leq n+1$, we have

$$c_i^+ \to \begin{cases} c_{k-i-1}^-, & 0 \leq i \leq k-1, \\ c_{i-k}^+, & k \leq i \leq n, \end{cases}$$

and

$$c_i^- \to \begin{cases} c_{i+k}^-, & 0 \le i \le n-k, \\ \\ c_{2n+1-i-k}^+, & n-k+1 \le i \le n. \end{cases}$$

Verify that as a result of the mapping $\theta \mapsto \theta - k\pi/(n+1)$, the *set* of $2n+2$ cotangents c_i^+ and c_i^- is mapped into itself.

To justify the above result for c_i^- with $i \ge n-k+1$, first show that

$$\cot(\phi + \tfrac{\pi}{2}) = \cot(\phi - \tfrac{\pi}{2}) = -\tan\phi.$$

Then write $\phi_i = \pi + \alpha_i$ and show that

$$c_i^- = \cot\frac{\theta - \pi - \alpha_i}{2} = \cot\frac{\theta + \pi - \alpha_i}{2} = c_{2n+1-i-k}^+.$$

Problem 2.5.4 Show that

$$\frac{1}{2^n}\frac{T_{n+1}(x)}{x - x_i^{(n)}} = \prod_{j \ne i}(x - x_j^{(n)}),$$

where the numbers $x_i^{(n)}$, $0 \le i \le n$, are the zeros of T_{n+1}. Put $x = \cos\theta$ and hence verify that

$$\frac{\cos(n+1)\theta}{\cos\theta - \cos\theta_i} = 2^n \prod_{j \ne i}(\cos\theta - \cos\theta_j),$$

where θ_i is defined in (2.162).

Problem 2.5.5 With $S_n = \sum_{i=1}^n 1/i$, deduce from the inequalities

$$\frac{1}{i+1} < \int_i^{i+1} \frac{dx}{x} < \frac{1}{i}, \qquad i \ge 1,$$

that

$$\int_1^{n+1} \frac{dx}{x} < S_n < 1 + \int_1^n \frac{dx}{x},$$

and hence

$$\log(n+1) < S_n < 1 + \log n, \qquad n \ge 1.$$

Problem 2.5.6 Verify that

$$\frac{(n+1)^2}{n(n+2)} = 1 + \frac{1}{n(n+2)}$$

decreases with n, and thus show that

$$\log(n+1) - \frac{1}{2}\log(n+2) = \frac{1}{2}\log n + \frac{1}{2}\log\frac{(n+1)^2}{n(n+2)} \le \frac{1}{2}\log n + \frac{1}{2}\log\frac{4}{3},$$

for all $n \ge 1$, with equality only for $n = 1$.

Problem 2.5.7 Show that $\Lambda_n(T)$, given by (2.169), may be expressed in the form

$$\Lambda_n(T) = \frac{1}{n+1}\sum_{i=0}^{n} \tan\frac{\theta_i}{2},$$

where θ_i is defined by (2.162).

2.6 The Modulus of Continuity

Definition 2.6.1 Given any function $f \in C[a,b]$, we define an associated function $\omega \in C[a,b]$ as

$$\omega(\delta) = \omega(f;[a,b];\delta) = \sup_{|x_1-x_2|\leq\delta} |f(x_1) - f(x_2)|. \tag{2.187}$$

We call $\omega(f;[a,b];\delta)$ the *modulus of continuity* of the function f. ■

We have written "sup" for *supremum*, meaning the least upper bound. We express the modulus of continuity in the simpler form $\omega(\delta)$ when it is clear which function f and interval $[a,b]$ are involved. It is not difficult to verify that $\omega \in C[a,b]$, given that $f \in C[a,b]$.

Example 2.6.1 For convenience, let us take the interval $[a,b]$ to be $[0,1]$. It is obvious from Definition 2.6.1 that

$$\omega(1;[0,1];\delta) = 0 \quad \text{and} \quad \omega(\delta;[0,1];\delta) = \delta.$$

To evaluate $\omega(\delta^2;\delta)$, let us write

$$x_1^2 - x_2^2 = (x_1 - x_2)(x_1 + x_2).$$

Thus, if $|x_1 - x_2| \leq \delta$, we have

$$|x_1^2 - x_2^2| \leq \delta|x_1 + x_2|,$$

and the right side of the above inequality is greatest when one of x_1, x_2 is 1 and the other is $1-\delta$. It is then clear that

$$\omega(\delta^2;[0,1];\delta) = \delta(2-\delta) = 1 - (1-\delta)^2.$$

More generally, it is not hard to see that

$$\omega(\delta^n;[0,1];\delta) = 1 - (1-\delta)^n$$

for all integers $n \geq 0$. ■

The above example helps give us some familiarity with the modulus of continuity, although the results obtained in it are of little intrinsic importance. It is not difficult to justify the following more substantial properties of the modulus of continuity.

Theorem 2.6.1 If $0 < \delta_1 \leq \delta_2$, then $\omega(\delta_1) \leq \omega(\delta_2)$. ∎

Theorem 2.6.2 A function f is uniformly continuous on the interval $[a, b]$ if and only if

$$\lim_{\delta \to 0} \omega(\delta) = 0. \qquad \blacksquare$$

We conclude this chapter with the statement of two important results due to Dunham Jackson (1888–1946) that express the minimax error for a function f in terms of moduli of continuity.

Theorem 2.6.3 Let $f \in C[-1, 1]$ and let

$$E_n(f) = \|f - p\|_\infty, \tag{2.188}$$

where $\|\cdot\|_\infty$ denotes the maximum norm on $[-1, 1]$, and $p \in P_n$ is the minimax approximation for $f \in C[a, b]$. Then

$$E_n(f) \leq 6\omega\left(\frac{1}{n}\right). \qquad \blacksquare \tag{2.189}$$

The second result of Jackson that we cite is applicable to functions that belong to $C^k[a, b]$, and gives an inequality that relates $E_n(f)$ to the modulus of continuity of the kth derivative of f.

Theorem 2.6.4 If $f \in C^k[-1, 1]$, then

$$E_n(f) \leq \frac{c}{n^k}\, \omega_k\left(\frac{1}{n-k}\right) \tag{2.190}$$

for $n > k$, where ω_k is the modulus of continuity of $f^{(k)}$, and

$$c = \frac{6^{k+1}e^k}{1+k}. \qquad \blacksquare$$

For proofs of these two theorems of Jackson, see Rivlin [48].

In addition to the modulus of continuity, there are other moduli that measure the "smoothness" of a function. These include moduli concerned with kth differences of a function. See the text by Sendov and Popov [51].

Problem 2.6.1 Show that

$$\omega(e^\delta; [0, 1]; \delta) = e - e^{1-\delta}.$$

Problem 2.6.2 Verify that for the sine function on the interval $[0, \pi/2]$, we have

$$\omega(\delta) = \sin \delta.$$

Find a class of functions f and intervals $[a, b]$ for which

$$\omega(f; [a, b]; \delta) = f(\delta).$$

Problem 2.6.3 If $|f'(x)| \leq M$ for $a \leq x \leq b$, show that

$$E_n(f) \leq \frac{6M}{n},$$

where $E_n(f)$ is defined in (2.188).

http://www.springer.com/978-0-387-00215-6

Interpolation and Approximation by Polynomials
Phillips, G.M.
2003, XIV, 312 p., Hardcover
ISBN: 978-0-387-00215-6