Chapter 5

Chebyshev Series

5.1 Introduction — Chebyshev series and other expansions

Many ways of expanding functions in infinite series have been studied. Indeed, the familiar Taylor series, Laurent series and Fourier series can all be regarded as expansions in functions orthogonal on appropriately chosen domains. Also, in the context of least-squares approximation, we introduced in Section 4.3.1 polynomial expansions whose partial sums coincide with best $\mathcal{L}_2$ approximations.

In the present chapter we link a number of these topics together in the context of expansions in Chebyshev polynomials (mainly of the first kind). Indeed a Chebyshev series is an important example of an orthogonal polynomial expansion, and may be transformed into a Fourier series or a Laurent series, according to whether the independent variable is real or complex. Such links are invaluable, not only in unifying mathematics but also in providing us with a variety of sources from which to obtain properties of Chebyshev series.

5.2 Some explicit Chebyshev series expansions

Defining an inner product $\langle f\, ,\, g\rangle$, as in Section 4.2, as

$$\langle f\, ,\, g\rangle = \int_{-1}^{1} w(x) f(x) g(x)\, \mathrm{d}x, \tag{5.1}$$

and restricting attention to the range $[-1, 1]$, the Chebyshev polynomials of first, second, third and fourth kinds are orthogonal with respect to the respective weight functions

$$w(x) = \frac{1}{\sqrt{1-x^2}},\ \sqrt{1-x^2},\ \sqrt{\frac{1+x}{1-x}} \text{ and } \sqrt{\frac{1-x}{1+x}}. \tag{5.2}$$

As we indicated in Section 4.3.1, the four kinds of Chebyshev series expansion of $f(x)$ have the form

$$f(x) \sim \sum_{i=0}^{\infty} c_i \phi_i(x) \tag{5.3}$$

where

$$c_i = \langle f\, ,\, \phi_i\rangle / \langle \phi_i\, ,\, \phi_i\rangle \tag{5.4}$$

and

$$\phi_i(x) = T_i(x),\ U_i(x),\ V_i(x) \text{ or } W_i(x) \tag{5.5}$$

corresponding to the four choices of weight function (5.2). Values for $\langle \phi_i\,,\,\phi_i \rangle$ were given in (4.11), (4.12), (4.13) and (4.14).

In the specific case of polynomials of the first kind, the expansion is

$$f(x) \sim \sum_{i=0}^{\infty}{}' c_i T_i(x) = \tfrac{1}{2} c_0 T_0(x) + c_1 T_1(x) + c_2 T_2(x) + \cdots \tag{5.6}$$

where

$$c_i = \frac{2}{\pi} \int_{-1}^{1} (1 - x^2)^{-\frac{1}{2}} f(x) T_i(x)\,\mathrm{d}x, \tag{5.7}$$

the dash, as usual, indicating that the first term in the series is halved. (Note the convenience in halving the first term, which enables us to use the same constant $2/\pi$ in (5.7) for every i including $i = 0$.)

There are several functions for which the coefficients c_i in (5.6) may be determined explicitly, although this is not possible in general.

EXAMPLE 5.1: Expansion of $f(x) = \sqrt{1 - x^2}$.

Here

$$\begin{aligned}
\frac{\pi}{2} c_i &= \int_{-1}^{1} T_i(x)\,\mathrm{d}x = \int_0^{\pi} \cos i\theta \sin\theta\,\mathrm{d}\theta \\
&= \tfrac{1}{2} \int_0^{\pi} [\sin(i+1)\theta - \sin(i-1)\theta]\,\mathrm{d}\theta \\
&= \tfrac{1}{2} \left[\frac{\cos(i-1)\theta}{i-1} - \frac{\cos(i+1)\theta}{i+1} \right]_0^{\pi} \quad (i \geq 1) \\
&= \tfrac{1}{2} \left(\frac{(-1)^{i-1} - 1}{i-1} - \frac{(-1)^{i+1} - 1}{i+1} \right)
\end{aligned}$$

and thus

$$c_{2k} = -\frac{4}{\pi(4k^2 - 1)}, \quad c_{2k-1} = 0 \quad (k = 1, 2, \ldots).$$

Also

$$c_0 = 4/\pi.$$

Hence,

$$\begin{aligned}
\sqrt{1 - x^2} &\sim -\frac{4}{\pi} \sum_{k=0}^{\infty}{}' \frac{T_{2k}(x)}{4k^2 - 1} \\
&= \frac{4}{\pi} \left(\tfrac{1}{2} T_0(x) - \tfrac{1}{3} T_2(x) - \tfrac{1}{15} T_4(x) - \tfrac{1}{35} T_6(x) - \cdots \right)
\end{aligned} \tag{5.8}$$

EXAMPLE 5.2: Expansion of $f(x) = \arccos x$.

This time,

$$\begin{aligned}
\frac{\pi}{2} c_i &= \int_{-1}^{1} (1-x^2)^{-\frac{1}{2}} \arccos x \, T_i(x) \, \mathrm{d}x \\
&= \int_0^{\pi} \theta \cos i\theta \, \mathrm{d}\theta \\
&= \left[\frac{\theta \sin i\theta}{i} \right]_0^{\pi} - \int_0^{\pi} \frac{\sin i\theta}{i} \, \mathrm{d}\theta \quad (i \geq 1) \\
&= \left[\frac{\theta \sin i\theta}{i} + \frac{\cos i\theta}{i^2} \right]_0^{\pi} \\
&= \frac{(-1)^i - 1}{i^2},
\end{aligned}$$

so that

$$c_{2k} = 0, \; c_{2k-1} = -\frac{2}{(2k-1)^2} \quad (k = 1, 2, \ldots).$$

Also

$$c_0 = \pi.$$

Hence,

$$\begin{aligned}
\arccos x &\sim \frac{\pi}{2} T_0(x) - \frac{4}{\pi} \sum_{k=1}^{\infty} \frac{T_{2k-1}(x)}{(2k-1)^2} \\
&= \frac{\pi}{2} T_0(x) - \frac{4}{\pi} \left(T_1(x) + \tfrac{1}{9} T_3(x) + \tfrac{1}{25} T_5(x) + \cdots \right) \qquad (5.9)
\end{aligned}$$

EXAMPLE 5.3: Expansion of $f(x) = \arcsin x$.

Here

$$\begin{aligned}
\frac{\pi}{2} c_i &= \int_{-1}^{1} (1-x^2)^{-\frac{1}{2}} \arcsin x \, T_i(x) \, \mathrm{d}x \\
&= \int_0^{\pi} \left(\frac{\pi}{2} - \theta \right) \cos i\theta \, \mathrm{d}\theta \\
&= \int_{-\pi/2}^{\pi/2} \phi \, \cos i \left(\frac{\pi}{2} - \phi \right) \, \mathrm{d}\phi.
\end{aligned}$$

Now

$$\begin{aligned}\frac{\pi}{2}c_{2k} &= \int_{-\pi/2}^{\pi/2} \phi\, \cos k(\pi - 2\phi)\, \mathrm{d}\phi \\ &= (-1)^k \int_{-\pi/2}^{\pi/2} \phi\, \cos 2k\phi\, \mathrm{d}\phi \\ &= 0\end{aligned}$$

(since the integrand is odd), while

$$\begin{aligned}\frac{\pi}{2}c_{2k-1} &= \int_{-\pi/2}^{\pi/2} \phi\, [\cos(k - \tfrac{1}{2})\pi\, \cos(2k-1)\phi + \sin(k-\tfrac{1}{2})\pi\, \sin(2k-1)\phi]\, \mathrm{d}\phi \\ &= 2(-1)^{k-1} \int_0^{\pi/2} \phi\, \sin(2k-1)\phi\, \mathrm{d}\phi \\ &= 2(-1)^{k-1} \left[-\frac{\phi\, \cos(2k+1)\phi}{2k-1} + \frac{\sin(2k-1)\phi}{(2k-1)^2} \right]_0^{\pi/2} \\ &= \frac{2}{(2k-1)^2}.\end{aligned}$$

Hence,

$$\arcsin x \sim \frac{4}{\pi} \sum_{k=1}^{\infty} \frac{T_{2k-1}(x)}{(2k-1)^2}. \tag{5.10}$$

Note that the expansions (5.9) and (5.10) are consistent with the relationship

$$\arccos x = \frac{\pi}{2} - \arcsin x.$$

This is reassuring! It is also clear that all three expansions (5.8)–(5.10) are uniformly convergent on $[-1, 1]$, since $|T_i(x)| \le 1$ and the expansions are bounded at worst by series which behave like the convergent series $\sum_1^{\infty} 1/k^2$. For example, the series (5.10) for $\arcsin x$ is bounded above and below by its values at ± 1, namely

$$\pm \frac{4}{\pi} \sum_{k=1}^{\infty} \frac{1}{(2k-1)^2}.$$

Since the series is uniformly convergent, the latter values must be $\pm\pi/2$.

The convergence of these examples must not, however, lead the reader to expect every Chebyshev expansion to be uniformly convergent; conditions for convergence are discussed later in this chapter.

To supplement the above examples, we list below a selection of other explicitly known Chebyshev expansions, with textbook references. Some of these examples will be set as exercises at the end of this chapter.

- From Rivlin (1974)

$$\operatorname{sgn} x \sim \frac{4}{\pi}\sum_{k=1}^{\infty}(-1)^{k-1}\frac{T_{2k-1}(x)}{2k-1}, \tag{5.11}$$

$$|x| \sim \frac{2}{\pi}T_0(x) + \frac{4}{\pi}\sum_{k=1}^{\infty}(-1)^{k-1}\frac{T_{2k}(x)}{4k^2-1}, \tag{5.12}$$

$$\frac{1}{a^2-x^2} \sim \frac{2}{a\sqrt{a^2-1}}\sum_{k=0}^{\infty}{}'(a-\sqrt{a^2-1})^{2k}T_{2k}(x) \ (a^2>1), \tag{5.13}$$

$$\frac{1}{x-a} \sim -\frac{2}{\sqrt{a^2-1}}\sum_{i=0}^{\infty}{}'(a-\sqrt{a^2-1})^{i}T_i(x) \quad (a>1). \tag{5.14}$$

- From Snyder (1966)

$$\arctan t \sim \frac{\pi}{8} + 2\sum_{k=0}^{\infty}(-1)^k\frac{v^{2k+1}}{2k+1}T_{2k+1}(x) \quad (t \text{ in } [0,1]) \tag{5.15}$$

$$\text{where } x = \frac{(\sqrt{2}+1)t-1}{(\sqrt{2}-1)t+1},\ v = \tan\frac{\pi}{16}, \tag{5.16}$$

$$\sin zx \sim 2\sum_{k=0}^{\infty}(-1)^k J_{2k+1}(z)T_{2k+1}(x) \tag{5.17}$$

where $J_k(z)$ is the Bessel function of the first kind,

$$e^{zx} \sim 2\sum_{k=0}^{\infty}{}' I_k(z)T_k(x) \tag{5.18}$$

$$\sinh zx \sim 2\sum_{k=0}^{\infty} I_{2k+1}(z)T_{2k+1}(x), \tag{5.19}$$

$$\cosh zx \sim 2\sum_{k=1}^{\infty}{}' I_{2k}(z)T_{2k}(x), \tag{5.20}$$

where $I_k(z)$ is the modified Bessel function of the first kind,

$$\frac{1}{1+x} \sim \sqrt{2}\sum_{i=0}^{\infty}{}'(-1)^i(3-2\sqrt{2})^iT_i^*(x) \quad (x \text{ in } [0,1]), \tag{5.21}$$

$$\ln(1+x) \sim \ln\left(\frac{3+2\sqrt{2}}{4}\right)T_0^*(x) + 2\sum_{i=1}^{\infty}(-1)^{i+1}\frac{(3-2\sqrt{2})^i}{i}T_i^*(x)$$
$$(x \text{ in } [0,1]), \tag{5.22}$$

$$\delta(x) \sim \frac{2}{\pi} \sideset{}{'}\sum_{i=0}^{\infty} (-1)^i T_{2i}(x) \tag{5.23}$$

where $\delta(x)$ is the 'Dirac delta function' with properties:

$$\delta(x) = 0 \text{ for } x \neq 0,$$

$$\int_{-\epsilon}^{\epsilon} \delta(x)\,\mathrm{d}x = 1 \text{ for } \epsilon > 0,$$

$$\int_{-1}^{1} \delta(x) f(x)\,\mathrm{d}x = f(0).$$

(The expansion (5.23) obviously cannot converge in any conventional sense.)

- From Fox & Parker (1968)

$$\frac{\arctan x}{x} \sim \sideset{}{'}\sum a_{2k} T_{2k}(x) \tag{5.24}$$

$$\text{where } a_{2k} = (-1)^k \sum_{s=k}^{\infty} 4 \frac{(\sqrt{2}-1)^{2s+1}}{2s+1}.$$

5.2.1 Generating functions

At least two well-known Chebyshev series expansions of functions involve a second variable (as did (5.17)–(5.20)), but in such a simple form (e.g., as a power of u) that they can be used (by equating coefficients) to generate formulae for the Chebyshev polynomials themselves. For this reason, such functions and their series are called *generating functions* for the Chebyshev polynomials.

- Our first generating function is given, by Snyder (1966) for example, in the form

$$F(u, z) = \mathrm{e}^{zu} \cos(u\sqrt{1-z^2}) = \sum_{n=0}^{\infty} \frac{u^n}{n!} T_n(z) \tag{5.25}$$

which follows immediately from the identity

$$\mathrm{Re}[\mathrm{e}^{u(\cos\theta + \mathrm{i}\sin\theta)}] = \sum_{n=0}^{\infty} \frac{u^n}{n!} \cos n\theta. \tag{5.26}$$

Although easily derived, (5.25) is not ideal for use as a generating function. The left-hand side expands into the product of two infinite series:

$$\sum_{n=0}^{\infty} \frac{u^n}{n!} T_n(z) = \mathrm{e}^{zu} \cos(u\sqrt{1-z^2}) = \sum_{i=0}^{\infty} \frac{z^i}{i!} u^i \sum_{j=0}^{\infty} \frac{(z^2-1)^j}{(2j)!} u^{2j}.$$

By equating coefficients of u^n, multiplying by $n!$ and simplifying, it is not difficult to derive the formula, previously quoted as (2.15) in Section 2.3.2,

$$T_n(z) = \sum_{k=0}^{\lfloor n/2 \rfloor} \left[(-1)^k \sum_{j=k}^{\lfloor n/2 \rfloor} \binom{n}{2j} \binom{j}{k} \right] z^{n-2k}, \tag{5.27}$$

where $\lfloor n/2 \rfloor$ denotes the integer part of $n/2$. However, although it is a compact expression, (5.27) is expensive to compute because of the double summation.

- A second and much more widely favoured generating function, given in Fox & Parker (1968), Rivlin (1974) and Snyder (1966), is

$$F(u,x) = \frac{1-ux}{1+u^2-2ux} = \sum_{n=0}^{\infty} T_n(x) u^n \quad (|u| < 1) \tag{5.28}$$

We follow the lead of Rivlin (1974) in favouring this. To obtain the coefficients in $T_n(x)$, we first note that

$$F(u, \tfrac{1}{2}x) = (1 - \tfrac{1}{2}ux) \frac{1}{1 - u(x-u)}, \tag{5.29}$$

and for any fixed x in $[-1, 1]$ the function $u(x-u)$ attains its greatest magnitude on $|u| \le \frac{1}{2}$ either at $u = \frac{1}{2}x$ (local maximum) or at one or other of $u = \pm\frac{1}{2}$. It follows that

$$-\tfrac{3}{4} \le u(x-u) \le \tfrac{1}{4} \quad (|u| \le \tfrac{1}{2},\ |x| \le 1)$$

and hence that the second factor in (5.29) can be expanded in a convergent series to give

$$\frac{1}{1-u(x-u)} = \sum_{n=0}^{\infty} u^n (x-u)^n = \sum_{n=0}^{\infty} c_n u^n, \text{ say,} \tag{5.30}$$

for $|u| \le \frac{1}{2}$. On equating coefficients of u^n in (5.30),

$$\begin{aligned} c_n &= x^n - \binom{n-1}{1} x^{n-2} + \binom{n-2}{2} x^{n-4} - \cdots + (-1)^k \binom{n-k}{k} x^{n-2k} + \\ &\quad + \cdots + (-1)^p \binom{n-p}{p} x^{n-2p} \end{aligned} \tag{5.31}$$

where $p = \lfloor n/2 \rfloor$. It is now straightforward to equate coefficients of u^n in (5.28), replacing x by $x/2$ and using (5.29)–(5.31), to obtain

$$T_n(x/2) = \sum_{k=0}^{\lfloor n/2 \rfloor} (-1)^k \left[\binom{n-k}{k} - \tfrac{1}{2} \binom{n-k-1}{k} \right] x^{n-2k} \tag{5.32}$$

where we interpret $\binom{n-k-1}{k}$ to be zero in case $n-k-1 < k$ (which arises when n is even and $k = p = n/2$). Since the polynomial equality (5.32) holds identically for $|x| \le 1$, it must hold for all x, so that we can in particular replace x by $2x$ to give

$$T_n(x) = \sum_{k=0}^{\lfloor n/2 \rfloor} (-1)^k 2^{n-2k-1} \left[2\binom{n-k}{k} - \binom{n-k-1}{k} \right] x^{n-2k}. \quad (5.33)$$

Simplifying this, we obtain finally

$$T_n(x) = \sum_{k=0}^{\lfloor n/2 \rfloor} (-1)^k 2^{n-2k-1} \frac{n}{(n-k)} \binom{n-k}{k} x^{n-2k} \quad (n > 0). \quad (5.34)$$

Formula (5.34) is essentially the same as formulae (2.16) and (2.18) of Section 2.3.2.

5.2.2 Approximate series expansions

The above special examples of explicit Chebyshev series generally correspond to cases where the integrals (5.4) can be evaluated mathematically. However, it is always possible to attempt to evaluate (5.4) numerically.

In the case of polynomials of the first kind, putting $x = \cos\theta$ in (5.7) gives

$$c_i = \frac{2}{\pi} \int_0^{\pi} f(\cos\theta) \cos i\theta \, \mathrm{d}\theta = \frac{1}{\pi} \int_0^{2\pi} f(\cos\theta) \cos i\theta \, \mathrm{d}\theta, \quad (5.35)$$

since the integrand is even and of period 2π in θ. The latter integral may be evaluated numerically by the trapezium rule based on any set of $2n+1$ points spaced at equal intervals of $h = \pi/n$, such as

$$\theta = \theta_k = \frac{(k - \frac{1}{2})\pi}{n}, \quad k = 1, 2, \ldots, 2n+1.$$

(With this choice, note that $\{\cos\theta_k\}$ are then the zeros of $T_n(x)$.) Thus

$$c_i = \frac{1}{\pi} \int_0^{2\pi} g_i(\theta) \, \mathrm{d}\theta = \frac{1}{\pi} \int_{\theta_1}^{\theta_{2n+1}} g_i(\theta) \, \mathrm{d}\theta \simeq \frac{h}{\pi} \sum_{k=1}^{2n+1}{}'' g_i(\theta_k), \quad (5.36)$$

where $g_i(\theta) := f(\cos\theta) \cos i\theta$ and where the double dash as usual indicates that the first and last terms of the summation are to be halved. But $g_i(\theta_1) = g_i(\theta_{2n+1})$, since g_i is periodic, and $g_i(2\pi - \theta) = g_i(\theta)$ so that $g_i(\theta_{2n+1-k}) = g_i(\theta_k)$. Hence (5.36) simplifies to

$$c_i \simeq \frac{2}{n} \sum_{k=1}^{n} g_i(\theta_k) = \frac{2}{n} \sum_{k=1}^{n} f(\cos\theta_k) \cos i\theta_k, \quad (i = 0, \ldots, n), \quad (5.37)$$

or, equivalently,

$$c_i \simeq \frac{2}{n}\sum_{k=1}^{n} f(x_k)T_i(x_k) \tag{5.38}$$

where $\{x_k\} = \{\cos\theta_k\}$ are the zeros of $T_n(x)$.

Formula (5.37) is what is commonly known as a *'discrete Fourier transform'*, and is a numerical approximation to the (continuous) Fourier transform (5.35). In fact, if the infinite expansion (5.6) is truncated after its first n terms (to give a polynomial of degree $(n-1)$), then the approximate series coefficients (5.37) yield the polynomial of degree $(k-1)$ which *exactly* interpolates $f(x)$ in the zeros $\{x_k\}$ of $T_n(x)$. So this approximate series method, based on efficient numerical quadrature, is really not a series method but an interpolation method. This assertion is proved and the 'Chebyshev interpolation polynomial' is discussed in depth in Chapter 6. The trapezium rule is a very accurate quadrature method for truly periodic trigonometric functions of θ, such as $g_i(\theta)$. Indeed, it is analogous to Gauss–Chebyshev quadrature for the original (x-variable) integral (5.7), which is known to be a very accurate numerical method (see Chapter 8). (On the other hand, the trapezium rule is a relatively crude method for the integration of non-trigonometric, non-periodic functions.) Hence, we can justifiably expect the Chebyshev interpolation polynomial to be a very close approximation to the partial sum (to the same degree) of the expansion (5.6). Indeed in practice these two approximations are virtually identical and to all intents and purposes interchangeable, as long as f is sufficiently smooth.

In Chapter 6, we shall state results that explicitly link the errors of a truncated Chebyshev series expansion and those of a Chebyshev interpolation polynomial. We shall also compare each of these in turn with the minimax polynomial approximation of the same degree. The interpolation polynomial will be discussed in this way in Chapter 6, but we give early attention to the truncated series expansion in Section 5.5 below.

5.3 Fourier–Chebyshev series and Fourier theory

Before we go any further, it is vital to link Chebyshev series to Fourier series, since this enables us to exploit a rich field as well as to simplify much of the discussion by putting it into the context of trigonometric functions. We first treat series of Chebyshev polynomials of the first kind, for which the theory is most powerful.

Suppose that $f(x)$ is square integrable ($\mathcal{L}_2$) on $[-1,1]$ with respect to the weight function $(1-x^2)^{-\frac{1}{2}}$, so that

$$\int_{-1}^{1} (1-x^2)^{-\frac{1}{2}} f(x)^2 \,\mathrm{d}x \tag{5.39}$$

is well defined (and finite). Now, with the usual change of variable, the function $f(x)$ defines a new function $g(\theta)$, where

$$g(\theta) = f(\cos\theta) \quad (0 \le \theta \le \pi). \tag{5.40}$$

We may easily extend this definition to all real θ by requiring that $g(\theta+2\pi) = g(\theta)$ and $g(-\theta) = g(\theta)$, when g becomes an even periodic function of period 2π. The integral (5.39) transforms into

$$\int_0^\pi g(\theta)^2 \, \mathrm{d}\theta,$$

so that g is $\mathcal{L}_2$-integrable with unit weight. Thus, g is ideally suited to expansion in a Fourier series.

The Fourier series of a general 2π-periodic function g may be written as

$$g(\theta) \sim \tfrac{1}{2}a_0 + \sum_{k=1}^{\infty}(a_k \cos k\theta + b_k \sin k\theta) \tag{5.41}$$

where

$$a_k = \frac{1}{\pi}\int_{-\pi}^{\pi} g(\theta)\cos k\theta \, \mathrm{d}\theta, \quad b_k = \frac{1}{\pi}\int_{-\pi}^{\pi} g(\theta)\sin k\theta \, \mathrm{d}\theta, \quad (k = 0, 1, 2, \ldots). \tag{5.42}$$

In the present case, since g is even in θ, all the b_k coefficients vanish, and the series simplifies to the Fourier cosine series

$$g(\theta) \sim {\sum_{k=0}^{\infty}}' a_k \cos k\theta \tag{5.43}$$

where

$$a_k = \frac{2}{\pi}\int_0^\pi g(\theta)\cos k\theta \, \mathrm{d}\theta. \tag{5.44}$$

If we now transform back to the x variable, we immediately deduce that

$$f(x) \sim {\sum_{k=0}^{\infty}}' a_k T_k(x) \tag{5.45}$$

where

$$a_k = \frac{2}{\pi}\int_{-1}^{1} (1-x^2)^{-\frac{1}{2}} f(x) T_k(x) \, \mathrm{d}x. \tag{5.46}$$

Thus, apart from the change of variables, the Chebyshev series expansion (5.45) is *identical* to the Fourier cosine series (5.43) and, indeed, the coefficients a_k occurring in the two expansions, derived from (5.44) and (5.46), have identical values.

5.3.1 $\mathcal{L}_2$-convergence

A fundamental property of the Fourier series of any $\mathcal{L}_2$-integrable function $g(\theta)$ is that it converges in the $\mathcal{L}_2$ norm. Writing the partial sum of order n of the Fourier expansion (5.41) as

$$(S_n^F g)(\theta) = \tfrac{1}{2}a_0 + \sum_{k=1}^{n}(a_k \cos k\theta + b_k \sin k\theta), \tag{5.47}$$

this means that

$$\left\|g - S_n^F g\right\|_2^2 = \int_{-\pi}^{\pi} [g(\theta) - (S_n^F g)(\theta)]^2 \,\mathrm{d}\theta \to 0 \text{ as } n \to \infty. \tag{5.48}$$

Lemma 5.1 *The partial sum (5.47) simplifies to*

$$(S_n^F g)(\theta) = \frac{1}{2\pi}\int_{-\pi}^{\pi} g(t+\theta)\frac{\sin(n+\frac{1}{2})t}{\sin\frac{1}{2}t}\,\mathrm{d}t = \frac{1}{2\pi}\int_{-\pi}^{\pi} g(t+\theta)W_n(\cos t)\,\mathrm{d}t, \tag{5.49}$$

where $W_n(x)$ is the Chebyshev polynomial of the fourth kind.

This is the classical *Dirichlet formula* for the partial Fourier sum.

Proof: It is easily shown that

$$\sideset{}{'}\sum_{k=0}^{n} \cos kt = \frac{1}{2}\frac{\sin(n+\frac{1}{2})t}{\sin\frac{1}{2}t}. \tag{5.50}$$

Substituting the expressions (5.42) for a_k and b_k in (5.47), we get

$$\begin{aligned}
(S_n^F g)(\theta) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} g(t)\,\mathrm{d}t + \frac{1}{\pi}\sum_{k=1}^{n}\int_{-\pi}^{\pi} g(t)(\cos kt\,\cos k\theta + \sin kt\,\sin k\theta)\,\mathrm{d}t \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} g(t)\,\mathrm{d}t + \frac{1}{\pi}\sum_{k=1}^{n}\int_{-\pi}^{\pi} g(t)\cos k(t-\theta)\,\mathrm{d}t \\
&= \frac{1}{\pi}\int_{-\pi}^{\pi} g(t)\sideset{}{'}\sum_{k=0}^{n}\cos k(t-\theta)\,\mathrm{d}t \\
&= \frac{1}{\pi}\int_{-\pi}^{\pi} g(t+\theta)\sideset{}{'}\sum_{k=0}^{n}\cos kt\,\mathrm{d}t \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} g(t+\theta)\frac{\sin(n+\frac{1}{2})t}{\sin\frac{1}{2}t}\,\mathrm{d}t \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} g(t+\theta)W_n(\cos t)\,\mathrm{d}t
\end{aligned}$$

as required. ●●

In the particular case of the function (5.40), which is even, the partial sum (5.47) simplifies to the partial sum of the Fourier cosine expansion

$$(S_n^F g)(\theta) = (S_n^{FC} g)(\theta) = \sum_{k=0}^{n}{}' a_k \cos k\theta. \tag{5.51}$$

This is identical, as we have said, to the partial sum of the Chebyshev series, which we write as

$$(S_n^T f)(x) = \sum_{k=0}^{n}{}' a_k T_k(x). \tag{5.52}$$

From (5.48) we immediately deduce, by changing variables, that

$$\int_{-1}^{1} (1-x^2)^{-\frac{1}{2}} [f(x) - (S_n^T f)(x)]^2 \, \mathrm{d}x \to 0 \text{ as } n \to \infty, \tag{5.53}$$

provided that $f(x)$ is $\mathcal{L}_2$ integrable on $[-1,1]$ with weight $(1-x^2)^{-\frac{1}{2}}$. Thus the Chebyshev series expansion is $\mathcal{L}_2$-convergent with respect to its weight function $(1-x^2)^{-\frac{1}{2}}$.

We know that the Chebyshev polynomials are mutually orthogonal on $[-1,1]$ with respect to the weight $(1-x^2)^{-\frac{1}{2}}$; this was an immediate consequence (see Section 4.2.2) of the orthogonality on $[0,\pi]$ of the cosine functions

$$\int_0^{\pi} \cos i\theta \cos j\theta \, \mathrm{d}\theta = 0 \quad (i \neq j).$$

Using the inner product

$$\langle f_1\,,\ f_2 \rangle := \int_{-1}^{1} (1-x^2)^{-\frac{1}{2}} f_1(x) f_2(x) \, \mathrm{d}x, \tag{5.54}$$

so that

$$a_k = \frac{2}{\pi} \langle T_k\,,\ f \rangle, \tag{5.55}$$

we find that

$$\begin{aligned}
\left\langle f - S_n^T f\,,\ f - S_n^T f \right\rangle &= \langle f\,,\ f \rangle - 2 \left\langle S_n^T f\,,\ f \right\rangle + \left\langle S_n^T f\,,\ S_n^T f \right\rangle \\
&= \|f\|^2 - 2 \sum_{k=0}^{n}{}' a_k \langle T_k\,,\ f \rangle + \tfrac{1}{4} a_0^2 \langle T_0\,,\ T_0 \rangle + \\
&\qquad + \sum_{k=1}^{n} a_k^2 \langle T_k\,,\ T_k \rangle \\
&\qquad \text{(from (5.52))} \\
&= \|f\|^2 - 2 \sum_{k=0}^{n}{}' a_k \frac{\pi}{2} a_k + \sum_{k=0}^{n}{}' a_k^2 \frac{\pi}{2}
\end{aligned}$$

(from (5.55) and (4.11))

$$= \|f\|^2 - \frac{\pi}{2} \sum_{k=0}^{n}{}' a_k^2.$$

From (5.53), this expression must tend to zero as $n \to \infty$. Therefore $\sum_{k=0}^{\prime\infty} a_k^2$ is convergent, and we obtain *Parseval's formula*:

$$\sum_{k=0}^{\infty}{}' a_k^2 = \frac{2}{\pi} \|f\|^2 = \frac{2}{\pi} \int_{-1}^{1} (1-x^2)^{-\frac{1}{2}} f(x)^2 \, \mathrm{d}x. \tag{5.56}$$

The following theorem summarises the main points above.

Theorem 5.2 *If $f(x)$ is $\mathcal{L}_2$-integrable with respect to the inner product (5.54), then its Chebyshev series expansion (5.45) converges in $\mathcal{L}_2$, according to (5.53). Moreover the infinite series $\sum_{k=0}^{\prime\infty} a_k^2$ is convergent to $2\pi^{-1} \|f\|^2$ (Parseval's formula).*

It is worthwhile at this juncture to insert a theorem on Fourier series, which, although weaker than the $\mathcal{L}_2$-convergence result, is surprisingly useful in its own right. We precede it with a famous inequality.

Lemma 5.3 (Hölder's inequality) *If $p \geq 1$, $q \geq 1$ and $1/p + 1/q = 1$, and if f is $\mathcal{L}_p$-integrable and g is $\mathcal{L}_q$-integrable over the same interval with the same weight, then*

$$\langle f\,,\, g \rangle \leq \|f\|_p \, \|g\|_q \,.$$

Proof: See, for instance, Hardy et al. (1952). ●●

From this lemma we may deduce the following.

Lemma 5.4 *If $1 \leq p_1 \leq p_2$ and f is $\mathcal{L}_{p_2}$-integrable over an interval, with respect to a (positive) weight $w(x)$ such that $\int w(x)\,\mathrm{d}x$ is finite, then f is $\mathcal{L}_{p_1}$-integrable with respect to the same weight, and*

$$\|f\|_{p_1} \leq C \, \|f\|_{p_2}$$

where C is a constant.

Proof: In Lemma 5.3, replace f by $|f|^{p_1}$, g by 1 and p by p_2/p_1, so that q is replaced by $p_2/(p_2 - p_1)$. This gives

$$\langle |f|^{p_1}\,,\, 1 \rangle \leq \left\| |f|^{p_1} \right\|_{p_2/p_1} \|1\|_{p_2/(p_2-p_1)}$$

or, written out in full,

$$\int w(x)\,|f(x)|^{p_1}\,\mathrm{d}x \le \left(\int w(x)\,|f(x)|^{p_2}\,\mathrm{d}x\right)^{p_1/p_2}\left(\int w(x)\,\mathrm{d}x\right)^{1-p_1/p_2}$$

and therefore, raising this to the power $1/p_1$,

$$\|f\|_{p_1} \le C\,\|f\|_{p_2}$$

where $C = \left(\int w(x)\,\mathrm{d}x\right)^{p_2-p_1}$. ●●

We can now state the theorem.

Theorem 5.5 *If $g(\theta)$ is $\mathcal{L}_2$-integrable on $[-\pi,\pi]$, then its Fourier series expansion converges in the $\mathcal{L}_1$ norm. That is:*

$$\int_{-\pi}^{\pi}\left|g(\theta)-(S_n^F g)(\theta)\right|\,\mathrm{d}\theta \to 0 \text{ as } n\to\infty.$$

Proof: By Lemma 5.4,

$$\left\|g - S_n^F g\right\|_1 \le C\left\|g - S_n^F g\right\|_2$$

with C a constant. Since a Fourier series converges in $\mathcal{L}_2$, the right-hand side tends to zero; hence, so does the left-hand side, and the result is proved. ●●

5.3.2 Pointwise and uniform convergence

So far, although we have established mean convergence for the Chebyshev series (4.24) in the sense of (5.53), this does not guarantee convergence at any particular point x, let alone ensuring uniform (i.e., $\mathcal{L}_\infty$) convergence. However, there are a number of established Fourier series results that we can use to ensure such convergence, either by making more severe assumptions about the function $f(x)$ or by modifying the way that we sum the Fourier series.

At the lowest level, it is well known that if $g(\theta)$ is continuous apart from a finite number of step discontinuities, then its Fourier series converges to g wherever g is continuous, and to the average of the left and right limiting values at each discontinuity. Translating this to $f(x)$, we see that if $f(x)$ is continuous in the interval $[-1,1]$ apart from a finite number of step discontinuities in the interior, then its Chebyshev series expansion converges to f wherever f is continuous, and to the average of the left and right limiting values at each discontinuity[1]. Assuming continuity everywhere, we obtain the following result.

[1] If g or f has finite step discontinuities, then a further problem is presented by the so-called *Gibbs phenomenon*: as the number of terms in the partial sums of the Fourier or Chebyshev series increases, one can find points approaching each discontinuity from either side where the error approaches a fixed non-zero value of around 9% of the height of the step, appearing to magnify the discontinuity.

Theorem 5.6 *If $f(x)$ is in $\mathcal{C}[-1,1]$, then its Chebyshev series expansion is pointwise convergent.*

To obtain *uniform* convergence of the Fourier series, a little more than continuity (and periodicity) is required of $g(\theta)$. A sufficient condition is that g should have bounded variation; in other words, that the absolute sum of all local variations (or oscillations) should not be unbounded. An alternative sufficient condition, which is neater but perhaps more complicated, is the *Dini–Lipschitz condition*:

$$\omega(\delta)\log\delta \to 0 \text{ as } \delta \to 0, \tag{5.57}$$

where $\omega(\delta)$ is a *modulus of continuity* for $g(\theta)$, such that

$$|g(\theta+\delta) - g(\theta)| \le \omega(\delta) \tag{5.58}$$

for all θ. The function $\omega(\delta)$ defines a level of continuity for g; for example, $\omega(\delta) = O(\delta)$ holds when g is differentiable, $\omega(\delta) \to 0$ implies only that g is continuous, while the Dini–Lipschitz condition lies somewhere in between. In fact, (5.57) assumes only 'infinitesimally more than continuity', compared with any assumption of differentiability. Translating the Fourier results to the x variable, we obtain the following.

Theorem 5.7 *If $f(x)$ is continuous and either of bounded variation or satisfying a Dini–Lipschitz condition on $[-1,1]$, then its Chebyshev series expansion is uniformly convergent.*

Proof: We need only show that bounded variation or the Dini–Lipschitz condition for $f(x)$ implies the same condition for $g(\theta) = f(\cos\theta)$. The bounded variation is almost obvious; Dini–Lipschitz follows from

$$\begin{aligned} |g(\theta+\delta) - g(\theta)| &= |f(\cos(\theta+\delta)) - f(\cos\theta)| \\ &\le \omega(\cos(\theta+\delta) - \cos\theta) \\ &\le \omega(\delta), \end{aligned}$$

since it is easily shown that $|\cos(\theta+\delta) - \cos\theta| \le |\delta|$ and that $\omega(\delta)$ is an increasing function of $|\delta|$. ●●

If a function is no more than barely continuous, then (Fejér 1904) we can derive uniformly convergent approximations from its Fourier expansion by averaging out the partial sums, and thus forming 'Cesàro sums' of the Fourier series

$$(\sigma_n^F g)(\theta) = \frac{1}{n}(S_0^F g + S_1^F g + \cdots + S_{n-1}^F g)(\theta). \tag{5.59}$$

Then $(\sigma_n^F g)(\theta)$ converges uniformly to $g(\theta)$ for every continuous function g. Translating this result into the Chebyshev context, we obtain not only uniformly convergent Chebyshev sums but also a famous corollary.

Theorem 5.8 *If $f(x)$ is continuous on $[-1, 1]$, then the Cesàro sums of its Chebyshev series expansion are uniformly convergent.*

Corollary 5.8A (Weierstrass's first theorem) *A continuous function may be arbitrarily well approximated on a finite interval in the minimax (uniform) sense by some polynomial of sufficiently high degree.*

Proof: This follows immediately from Theorem 5.8, since $(\sigma_n^T f)(x)$ is a polynomial of degree n which converges uniformly to $f(x)$ as $n \to \infty$. ●●

5.3.3 Bivariate and multivariate Chebyshev series expansions

Fourier and Chebyshev series are readily extended to two or more variables by tensor product techniques. Hobson (1926, pages 702–710) gives an early and unusually detailed discussion of the two-dimensional Fourier case and its convergence properties, and Mason (1967) was able to deduce (by the usual $x = \cos\theta$ transformation) a convergence result for double Chebyshev series of the first kind. This result is based on a two-dimensional version of 'bounded variation' defined as follows.

Definition 5.1 *Let $f(x, y)$ be defined on $D := \{-1 \le x \le 1; -1 \le y \le 1\}$; let $\{x_r\}$ and $\{y_r\}$ denote monotone non-decreasing sequences of $n+1$ values with $x_0 = y_0 = -1$ and $x_n = y_n = +1$; let*

$$\Sigma_1 := \sum_{r=1}^{n} |f(x_r, y_r) - f(x_{r-1} - y_{r-1})| ,$$

$$\Sigma_2 := \sum_{r=1}^{n} |f(x_r, y_{n-r+1}) - f(x_{r-1} - y_{n-r})| .$$

Then $f(x, y)$ is of bounded variation *on D if Σ_1 and Σ_2 are bounded for all possible sequences $\{x_r\}$ and $\{y_r\}$ and for every $n > 0$.*

Theorem 5.9 *If $f(x, y)$ is continuous and of bounded variation in*

$$S : \{-1 \le x \le 1; -1 \le y \le 1\},$$

and if one of its partial derivatives is bounded in S, then f has a double Chebyshev expansion, uniformly convergent on S, of the form

$$f(x, y) = {\sum_{i=0}^{\infty}}' {\sum_{j=0}^{\infty}}' a_{ij} T_i(x) T_j(y).$$

However, Theorem 5.9, based on bounded variation, is not a natural extension of Theorem 5.7, and it happens that the use of the Dini–Lipschitz condition is much easier to generalise.

There are detailed discussions by Mason (1980, 1982) of multivariate Chebyshev series, interpolation, expansion and near-best approximation formulae, with Lebesgue constants and convergence properties. The results are generally exactly what one would expect; for example, multivariate Lebesgue constants are products of univariate Lebesgue constants. Convergence, however, is a little different, as the following result illustrates.

Theorem 5.10 (Mason 1978) *If $f(x_1, \ldots, x_N)$ satisfies a Lipschitz condition of the form*

$$\sum_{j=1}^{N} \omega_j(\delta_j) \prod_{j=1}^{N} \log \delta_j \to 0 \text{ as } \delta_j \to 0,$$

where $\omega_j(\delta_j)$ is the modulus of continuity of f in the variable x_j, then the multivariate Fourier series of f, the multivariate Chebyshev series of f and the multivariate polynomial interpolating f at a tensor product of Chebyshev zeros all converge uniformly to f as $n_j \to \infty$. (In the case of the Fourier series, f must also be periodic for convergence on the whole hypercube.)

Proof: The proof employs two results: that the uniform error is bounded by

$$C \sum_j \omega_j \left(\frac{1}{n_j + 1} \right)$$

(Handscomb 1966, Timan 1963, Section 5.3) and that the Lebesgue constant is of order $\prod \log(n_j + 1)$. ●●

5.4 Projections and near-best approximations

In the previous section, we denoted a Chebyshev series partial sum by $S_n^T f$, the symbol S_n^T being implicitly used to denote an operator applied to f. In fact, the operator in question belongs to an important family of operators, which we term *projections*, which has powerful properties. In particular, we are able to estimate how far any projection of f is from a best approximation to f in any given norm.

Definition 5.2 *A* projection *P, mapping elements of a vector space $\mathcal{F}$ onto elements of a subspace $\mathcal{A}$ of $\mathcal{F}$, has the following properties:*

1. *P is a bounded operator;*

   i.e., there is a finite constant C such that $\|Pf\| \le C \|f\|$ for all f in $\mathcal{F}$;

2. *P is a linear operator;*

   i.e., $P(\lambda_1 f_1 + \lambda_2 f_2) = \lambda_1 P f_1 + \lambda_2 P f_2$, where λ_1, λ_2 are scalars and f_1, f_2 are in $\mathcal{F}$;

3. *P is idempotent;*

   i.e., $P(Pf) = Pf$ for all f in $\mathcal{F}$.

   Another way of expressing this, writing $P^2 f$ for $P(Pf)$, is to say that

$$P^2 = P. \tag{5.60}$$

The last property is a key one, ensuring that elements of the subspace $\mathcal{A}$ are invariant under the operator P. This is readily deduced by noting that, for any g in $\mathcal{A}$, there are elements f in $\mathcal{F}$ such that $g = Pf$, and hence

$$Pg = P(Pf) = Pf = g.$$

The mapping S_n^T of $\mathcal{C}[-1, 1]$ onto the space Π_n of polynomials of degree n is clearly a projection. (We leave the verification of this as an exercise to the reader.) In particular, it is clear that S_n^T is idempotent, since the Chebyshev partial sum of degree n of a polynomial of degree n is clearly that same polynomial.

On the other hand, not all approximation operators are projections. For example, the Cesàro sum operator defined in (5.47) is not idempotent, since in averaging the partial sums it alters the (trigonometric) polynomial. Also the minimax approximation operator B_n from $\mathcal{C}[-1, 1]$ onto a subspace $\mathcal{A}_n$ is nonlinear, since the minimax approximation to $\lambda_1 f_1 + \lambda_2 f_2$ is not in general $\lambda_1 B_n f_1 + \lambda_2 B_n f_2$. However, if we change to the L_2 norm, then the best approximation operator does become a projection, since it is identical to the partial sum of an orthogonal expansion.

Since we shall go into detail about the subject of near-best approximations, projections and minimal projections in a later chapter (Chapter 7), we restrict discussion here to general principles and to Chebyshev series (and related Fourier series) partial sum projections. In particular, we concentrate on L_∞ approximation by Chebyshev series of the first kind.

How then do we link projections to best approximations? The key to this is the fact that any projection (in the same vector space setting) takes a best approximation into itself. Consider in particular the setting

$$\mathcal{F} = \mathcal{C}[-1, 1], \quad \mathcal{A} = \Pi_n = \{\text{polynomials of degree} \le n\} \subset \mathcal{F}.$$

Now suppose that P_n is any projection from $\mathcal{F}$ onto $\mathcal{A}$ and that B_n is the best approximation operator in a given norm $\|\cdot\|$, and let I denote the identity operator. Then the best polynomial approximation of degree n to any f in $\mathcal{F}$ is $B_n f$ and, since this is invariant under P_n,

$$(I - P_n)(B_n f) = B_n f - P_n(B_n f) = B_n f - B_n f = 0. \tag{5.61}$$

The error in the approximation $P_n f$, which we wish to compare with the error in $B_n f$, is therefore given by

$$f - P_n f = (I - P_n)f = (I - P_n)f - (I - P_n)(B_n f) = (I - P_n)(f - B_n f), \quad (5.62)$$

using the fact that I, P_n and hence $(I - P_n)$ are linear. (Indeed, $(I - P_n)$ is another projection, since $(I - P_n)^2 = I - 2P_n + P_n^2 = I - P_n$, so that $(I - P_n)$ is also idempotent.)

In order to go further, we need to define the norm of a linear operator, which we do in precisely the same way as the norm of a matrix. We also need to be able to split up the norm of an operator applied to a function.

Definition 5.3 (Operator norm) *If T is a linear operator from a normed linear space into itself, or into another normed linear space, then the* operator norm $\|T\|$ *of T is defined to be the upper bound (if it exists)*

$$\|T\| = \sup_{f \neq 0} \frac{\|Tf\|}{\|f\|} \quad (5.63)$$

or, equivalently,

$$\|T\| = \sup_{\|f\|=1} \|Tf\| . \quad (5.64)$$

Lemma 5.11

$$\|Tf\| \leq \|T\| \, \|f\| . \quad (5.65)$$

Proof: Clearly $\|T\| \geq \|Tf\| / \|f\|$ for any particular f, since $\|T\|$ is the supremum over all f by the definition (5.63). ●●

We may now deduce the required connection between $P_n f$ and $B_n f$.

Theorem 5.12 *For P_n and B_n defined as above, operating from $\mathcal{F}$ onto $\mathcal{A}$,*

$$\|f - P_n f\| \leq \|I - P_n\| \, \|f - B_n f\| , \quad (5.66a)$$

$$\|f - P_n f\| \leq (1 + \|P_n\|) \, \|f - B_n f\| . \quad (5.66b)$$

Proof: Inequality (5.66a) follows immediately from (5.62) and (5.65).

Inequality (5.66b) then follows immediately from the deduction that for linear operators P and Q from $\mathcal{F}$ onto $\mathcal{A}$

$$\begin{aligned}
\|P + Q\| &= \sup_{\|f\|=1} \|(P + Q)f\| \\
&= \sup_{\|f\|=1} \|Pf + Qf\| \\
&\leq \sup_{\|f\|=1} (\|Pf\| + \|Qf\|) \\
&\leq \sup_{\|f\|=1} \|Pf\| + \sup_{\|f\|=1} \|Qf\| \\
&= \|P\| + \|Q\| .
\end{aligned}$$

Hence

$$\|I - P_n\| \leq \|I\| + \|P_n\| = 1 + \|P_n\| . \quad \bullet\bullet$$

Both formulae (5.66a) and (5.66b) in Theorem 5.12 give bounds on the error $\|f - P_n f\|$ in terms of absolute magnification factors $\|I - P_n\|$ or $(1 + \|P_n\|)$ on the best error $\|f - B_n f\|$. Clearly minimisation of these factors is a way of providing the best bound possible in this context. In particular Cheney & Price (1970) give the following definitions.

Definition 5.4 (Minimal projection) *A* minimal projection *is a projection P_n from $\mathcal{F}$ onto $\mathcal{A}$ for which $\|P_n\|$ (and hence $(1 + \|P_n\|)$) is as small as possible.*

Definition 5.5 (Cominimal projection) *A* cominimal projection *is a projection P_n from $\mathcal{F}$ onto $\mathcal{A}$ for which $\|I - P_n\|$ is as small as possible.*

Sometimes we are able to establish that a given projection is minimal (or cominimal) — examples of minimal projections include (in appropriate settings) the partial sums of Fourier, Taylor and Laurent series. However, even if a projection is not minimal, the estimates (5.66a) and (5.66b) are very useful. In particular, from (5.66b), the value of $\|P_n\|$ provides a bound on the relative closeness of the error in the approximation $P_n f$ to the error of the best approximation. Mason (1970) quantified this idea in practical terms by introducing a specific definition of a 'near-best approximation', reproduced here from Definition 3.2 of Chapter 3.

Definition 5.6 (Near-best and near-minimax approximations) *An approximation $f_N^*(x)$ in $\mathcal{A}$ is said to be* near-best within a relative distance ρ *if*

$$\|f - f_N^*\| \leq (1 + \rho) \|f - f_B^*\| ,$$

where ρ is a specified positive scalar and $f_B^*(x)$ is a best approximation. In the case of the L_∞ (minimax) norm, such an f^* is said to be near-minimax *within a relative distance ρ.*

Lemma 5.13 *If P_n is a projection of $\mathcal{F}$ onto $\mathcal{A} \subset \mathcal{F}$, and f is an element of $\mathcal{F}$ then, as an approximation to f, $P_n f$ is near-best within a relative distance $\|P_n\|$.*

Proof: This follows immediately from (5.66b). ●●

The machinery is now available for us to attempt to quantify the closeness of a Fourier–Chebyshev series partial sum to a minimax approximation. The aim is to bound or evaluate $\|P_n\|$, and this is typically achieved by first finding a formula for $P_n f$ in terms of f.

5.5 Near-minimax approximation by a Chebyshev series

Consider a function $f(x)$ in $\mathcal{F} = \mathcal{C}[-1,1]$ (i.e., a function continuous for $-1 \le x \le 1$) which has a Chebyshev partial sum of degree n of the form

$$(S_n^T f)(x) = {\sum_{k=0}^{n}}{}' c_k T_k(x), \quad c_k = \frac{2}{\pi}\int_{-1}^{1} \frac{f(x)T_k(x)}{\sqrt{1-x^2}}\,\mathrm{d}x. \tag{5.67}$$

If, as in Section 5.3, we define

$$g(\theta) = f(\cos\theta)$$

then we obtain the equivalent Fourier cosine series partial sum

$$(S_n^{FC} g)(\theta) = {\sum_{k=0}^{n}}{}' c_k \cos k\theta, \quad c_k = \frac{2}{\pi}\int_{0}^{\pi} g(\theta)\cos k\theta\,\mathrm{d}\theta. \tag{5.68}$$

The operator S_n^{FC} can be identified as the restriction of the Fourier projection S_n^F to the space $\mathcal{C}^0_{2\pi,\mathrm{e}}$ of continuous functions that are both periodic of period 2π and even. Indeed, there is a one-to-one relation between f in $\mathcal{C}[-1,1]$ and g in $\mathcal{C}^0_{2\pi,\mathrm{e}}$ under the mapping $x = \cos\theta$, in which each term of the Chebyshev series of f is related to the corresponding term of the Fourier cosine series of g.

Now, from Lemma 5.1, we know that S_n^F may be expressed in the integral form (5.49)

$$(S_n^F g)(\theta) = \frac{1}{2\pi}\int_{-\pi}^{\pi} g(t+\theta)\frac{\sin(n+\frac{1}{2})t}{\sin\frac{1}{2}t}\,\mathrm{d}t. \tag{5.69}$$

From this expression, bounding g by its largest absolute value, we get the inequality

$$\left|(S_n^F g)(\theta)\right| \le \lambda_n \left\|g\right\|_\infty \tag{5.70}$$

where

$$\begin{aligned} \lambda_n &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\left|\frac{\sin(n+\frac{1}{2})t}{\sin\frac{1}{2}t}\right|\mathrm{d}t = \\ &= \frac{1}{\pi}\int_{0}^{\pi}\left|\frac{\sin(n+\frac{1}{2})t}{\sin\frac{1}{2}t}\right|\mathrm{d}t \; \left[= \frac{1}{\pi}\int_{-1}^{1}\frac{|W_n(x)|}{\sqrt{1-x^2}}\,\mathrm{d}x\right]. \end{aligned} \tag{5.71}$$

Taking the supremum over θ of the left-hand side of (5.70),

$$\left\|S_n^F g\right\|_\infty \le \lambda_n \left\|g\right\|_\infty, \tag{5.72}$$

whence from (5.63) it follows that

$$\left\|S_n^F\right\|_\infty \le \lambda_n \tag{5.73}$$

and, *a fortiori*, since

$$\sup_{g\in\mathcal{C}^0_{2\pi,\mathrm{e}}} \frac{\left\|S_n^{FC} g\right\|_\infty}{\|g\|_\infty} = \sup_{g\in\mathcal{C}^0_{2\pi,\mathrm{e}}} \frac{\left\|S_n^{F} g\right\|_\infty}{\|g\|_\infty} \leq \sup_{g\in\mathcal{C}^0_{2\pi}} \frac{\left\|S_n^{F} g\right\|_\infty}{\|g\|_\infty},$$

that

$$\left\|S_n^{FC}\right\|_\infty \leq \left\|S_n^{F}\right\|_\infty \leq \lambda_n. \tag{5.74}$$

As a consequence of the one-to-one relationship between every $f(x)$ in $\mathcal{C}[-1,1]$ and a corresponding $g(\theta)$ in $\mathcal{C}^0_{2\pi,\mathrm{e}}$, it also immediately follows that

$$\left\|S_n^{T}\right\|_\infty = \left\|S_n^{FC}\right\|_\infty \leq \lambda_n \quad \text{(on the space } \mathcal{C}[-1,1]). \tag{5.75}$$

From Theorem 5.12 we may therefore assert that $(S_n^T f)(x)$ is near-minimax within a relative distance λ_n. So, how small or large is λ_n? Or, in other words, have we obtained a result that is really useful? The answer is rather interesting.

The constant λ_n is a famous one, the *Lebesgue constant*, and it is not difficult to show that

$$\lambda_n > \frac{4}{\pi^2}\log n. \tag{5.76}$$

So λ_n tends to infinity with n, which seems at first discouraging. However, $\log n$ grows extremely slowly, and indeed λ_n does not exceed 4 for $n \leq 500$. Thus, although it is true to say that $S_n^T f$ becomes relatively further away (without bound) from the best approximation $B_n f$ as n increases, it is also true to say that for all practical purposes $S_n^T f$ may be correctly described as a near-minimax approximation. Some values of λ_n are given in Table 5.1.

Table 5.1: Values of the Lebesgue constant

n	λ_n	n	λ_n	n	λ_n
1	1.436	10	2.223	100	3.139
2	1.642	20	2.494	200	3.419
3	1.778	30	2.656	300	3.583
4	1.880	40	2.770	400	3.699
5	1.961	50	2.860	500	3.789

More precise estimates than (5.76) have been derived by a succession of authors; for instance, Cheney & Price (1970) give the asymptotic formula

$$\lambda_n = \frac{4}{\pi^2}\log n + 1.2703\ldots + O(1/n). \tag{5.77}$$

5.5.1 Equality of the norm to λ_n

We have not yet fully completed the above analysis, since it turns out in fact that we may replace '$\le$' by '$=$' in (5.73), (5.74) and (5.75). This does not improve the practical observations above, but it does tell us that we cannot find a better bound than that given by (5.71).

To establish equality, it suffices to show that one particular function $g(\theta)$ exists in $\mathcal{C}^0_{2\pi,\mathrm{e}}$, and one value of θ exists in $[0, \pi]$, for which

$$\left|(S_n^F g)(\theta)\right| > \lambda_n \left\|g\right\|_\infty - \delta \tag{5.78}$$

with δ arbitrarily small — for then we must have

$$\left\|S_n^{FC} g\right\|_\infty = \left\|S_n^F g\right\|_\infty \ge \lambda_n \left\|g\right\|_\infty \tag{5.79}$$

and hence, from (5.72),

$$\left\|S_n^{FC} g\right\|_\infty = \left\|S_n^F g\right\|_\infty = \lambda_n \left\|g\right\|_\infty \tag{5.80}$$

and finally

$$\left\|S_n^T\right\|_\infty = \left\|S_n^{FC}\right\|_\infty = \left\|S_n^F\right\|_\infty = \lambda_n. \tag{5.81}$$

Proof: First, define

$$g_D(\theta) := \operatorname{sgn}\left(\frac{\sin(n+\frac{1}{2})\theta}{\sin\frac{1}{2}\theta}\right). \tag{5.82}$$

where

$$\operatorname{sgn}(x) := \begin{cases} +1, & x > 0 \\ 0, & x = 0 \\ -1, & x < 0. \end{cases}$$

Then

$$\left\|g_D\right\|_\infty = 1. \tag{5.83}$$

Moreover g_D is continuous apart from a finite number of step discontinuities, and is an even periodic function of period 2π. It is now a technical matter, which we leave as an exercise to the reader (Problem 6), to show that it is possible to find a continuous function $g_C(\theta)$, which also is even and periodic, such that

$$\left\|g_C - g_D\right\|_1 := \int_{-\pi}^{\pi} \left|g_C(t) - g_D(t)\right| \,\mathrm{d}t < \epsilon$$

and such that $\left\|g_C\right\|_\infty$ is within ϵ of unity, where ϵ is a specified small quantity.

Then, noting that n is fixed and taking θ as 0 in (5.69)

$$\begin{aligned}
(S_n^F g_C)(0) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} g_C(t)\frac{\sin(n+\frac{1}{2})t}{\sin\frac{1}{2}t}\,\mathrm{d}t \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} (g_C(t) - g_D(t))\frac{\sin(n+\frac{1}{2})t}{\sin\frac{1}{2}t}\,\mathrm{d}t + \frac{1}{2\pi}\int_{-\pi}^{\pi} g_D(t)\frac{\sin(n+\frac{1}{2})t}{\sin\frac{1}{2}t}\,\mathrm{d}t \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} (g_C(t) - g_D(t))\frac{\sin(n+\frac{1}{2})t}{\sin\frac{1}{2}t}\,\mathrm{d}t + \lambda_n, \text{ from (5.71),}
\end{aligned}$$

while

$$\left|\frac{1}{2\pi}\int_{-\pi}^{\pi}(g_C(t)-g_D(t))\frac{\sin(n+\frac{1}{2})t}{\sin\frac{1}{2}t}\,\mathrm{d}t\right| \le \frac{1}{2\pi}\left\|\frac{\sin(n+\frac{1}{2})t}{\sin\frac{1}{2}t}\right\|_\infty \|g_C-g_D\|_1$$
$$= \frac{1}{2\pi}\|W_n\|_\infty \|g_C-g_D\|_1$$
$$\le \frac{1}{\pi}(2n+1)\epsilon$$

since $|W_n(x)|$ has a greatest value of $2n+1$ (attained at $x=1$).

Thus

$$\left|(S_n^F g_C)(0)\right| \ge \lambda_n - \frac{1}{\pi}(2n+1)\epsilon$$

and

$$\lambda_n \|g_C\|_\infty \le \lambda_n(1+\epsilon).$$

For any small δ, we can then make ϵ so small that (5.78) is satisfied at $\theta = 0$ by $g = g_C$. ●●

5.6 Comparison of Chebyshev and other orthogonal polynomial expansions

The partial sum (5.47) of a Fourier series represents a projection from the space $\mathcal{C}_{2\pi}^0$ onto the corresponding subspace of sums of sine and cosine functions, that is both minimal and cominimal (in the minimax norm). This may be shown (Cheney 1966, Chapter 6) by considering any other projection operator P from $\mathcal{C}_{2\pi}^0$ onto the space of linear combinations of sines and cosines up to $\cos n\theta$ and $\sin n\theta$, letting T_λ be the shift operator defined by

$$(T_\lambda f)(\theta) = f(\theta+\lambda) \text{ for all } \theta$$

and showing that

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}(T_{-\lambda}PT_\lambda f)(\theta)\,\mathrm{d}\lambda \equiv (S_n^F f)(\theta). \tag{5.84}$$

Since $\|T_\lambda\|_\infty = \|T_{-\lambda}\|_\infty = 1$, we can then deduce that

$$\|P\|_\infty \ge \left\|S_n^F\right\|_\infty,$$

so that S_n^F is minimal. It follows likewise, since (5.84) implies

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}(T_{-\lambda}(I-P)T_\lambda f)(\theta)\,\mathrm{d}\lambda \equiv ((I-S_n^F)f)(\theta), \tag{5.85}$$

that S_n^F is cominimal.

Thus we can say that the partial sums of the Fourier expansion of a continuous periodic function 'converge faster', in terms of their minimax error bounds, than any other approximations obtained by projection onto subspaces of trigonometric polynomials.

Remembering what we have successfully done on many previous occasions, we might have supposed that, by means of the substitution $x = \cos\theta$, we could have derived from the above a proof of an analogous conjecture that the partial sums of a first-kind Chebyshev expansion of a continuous function on $[-1, 1]$ converge faster than any other polynomial approximations obtained by projection; that is, than the partial sums of an expansion in polynomials orthogonal with respect to any other weight. Unfortunately, this is not possible — to do so we should first have needed to show that S_n^F was minimal and cominimal on the space $\mathcal{C}^0_{2\pi,\mathrm{e}}$ of even periodic functions, but the above argument then breaks down since the shift operator T_λ does not in general transform an even function into an even function.

The conjecture closely reflects practical experience, nevertheless, so that a number of attempts have been made to justify it.

In order to show first-kind Chebyshev expansions to be superior to expansions in other ultraspherical polynomials, Lanczos (1952) argued as follows:

Proof: The expansion of a function $f(x)$ in ultraspherical polynomials is

$$f(x) = \sum_{k=0}^{\infty} c_k^{(\alpha)} P_k^{(\alpha)}(x) \tag{5.86}$$

with coefficients given by

$$c_k^{(\alpha)} = \frac{\displaystyle\int_{-1}^{1} (1-x^2)^\alpha f(x) P_k^{(\alpha)}(x)\,\mathrm{d}x}{\displaystyle\int_{-1}^{1} (1-x^2)^\alpha [P_k^{(\alpha)}(x)]^2\,\mathrm{d}x}. \tag{5.87}$$

Using the Rodrigues formula (4.29), this gives us

$$c_k^{(\alpha)} = \frac{\displaystyle\int_{-1}^{1} f(x) \frac{\mathrm{d}^k}{\mathrm{d}x^k}(1-x^2)^{k+\alpha}\,\mathrm{d}x}{\displaystyle\int_{-1}^{1} P_k^{(\alpha)}(x) \frac{\mathrm{d}^k}{\mathrm{d}x^k}(1-x^2)^{k+\alpha}\,\mathrm{d}x} \tag{5.88}$$

or, integrating k times by parts,

$$c_k^{(\alpha)} = \frac{\displaystyle\int_{-1}^{1} \frac{\mathrm{d}^k}{\mathrm{d}x^k} f(x)\,(1-x^2)^{k+\alpha}\,\mathrm{d}x}{\displaystyle\int_{-1}^{1} \frac{\mathrm{d}^k}{\mathrm{d}x^k} P_k^{(\alpha)}(x)\,(1-x^2)^{k+\alpha}\,\mathrm{d}x}$$

$$= \frac{\displaystyle\int_{-1}^{1} f^{(k)}(x)\,(1-x^2)^{k+\alpha}\,\mathrm{d}x}{\displaystyle k!K_k^{(\alpha)}\int_{-1}^{1}(1-x^2)^{k+\alpha}\,\mathrm{d}x}, \tag{5.89}$$

where $K_k^{(\alpha)}$ is the coefficient of the leading power x^k in $P_k^{(\alpha)}(x)$.

As $k \to \infty$, then claims Lanczos, the factor $(1-x^2)^{k+\alpha}$ in each integrand approaches a multiple of the delta function $\delta(x)$, so that

$$c_k^{(\alpha)} \sim \frac{f^{(k)}(0)}{k!K_k^{(\alpha)}}. \tag{5.90}$$

Since we have not yet specified a normalisation for the ultraspherical polynomials, we may take them all to be monic polynomials ($K_k^{(\alpha)} = 1$), so that in particular $P_k^{(-\frac{1}{2})}(x) = 2^{1-k}T_k(x)$. Then the minimax norm of the kth term of the expansion (5.86) is given by

$$\left|c_k^{(\alpha)}\right| \left\|P_k^{(\alpha)}\right\|_\infty \sim \left|\frac{f^{(k)}(0)}{k!}\right| \left\|P_k^{(\alpha)}\right\|_\infty. \tag{5.91}$$

But (Corollary 3.4B) $P_k^{(-\frac{1}{2})}(x) = 2^{1-k}T_k(x)$ is the monic polynomial of degree k with smallest minimax norm on $[-1,1]$. Hence the terms of the first-kind Chebyshev expansion are in the limit smaller in minimax norm, term by term, than those of any other ultraspherical expansion. ●●

This argument is not watertight. First, it assumes that $f^{(k)}(0)$ exists for all k. More seriously, it assumes that these derivatives do not increase too rapidly with k — otherwise the asymptotic form (5.90) cannot be justified. By use of formulae expressing the ultraspherical polynomials as linear combinations of Chebyshev polynomials, and by defining a somewhat contrived measure of the rate of convergence, Handscomb (1973) was able to find a sense in which the first-kind Chebyshev expansion converges better than ultraspherical expansions with $\alpha > -\frac{1}{2}$, but was unable to extend this at all satisfactorily to the case where $-1 < \alpha < -\frac{1}{2}$. Subsequently, Light (1978) computed the norms of a number of ultraspherical projection operators, finding that they all increased monotonically with α, so that the Chebyshev projection cannot be minimal. However, this did not answer the more important question of whether the Chebyshev projection is cominimal.

Later again, Light (1979) proved, among other results, that the first-kind Chebyshev expansion of a function f converges better than ultraspherical expansions with $\alpha > -\frac{1}{2}$, in the conventional sense that

$$\left\|f - S_n^T f\right\|_\infty < \left\|f - \sum_{k=0}^{n} c_k^{(\alpha)} P_k^{(\alpha)}\right\|_\infty \quad \text{for sufficiently large } n, \tag{5.92}$$

provided that f has a Chebyshev expansion $\sum_k b_k T_k$ with

$$2^k\,|b_k| \to A \text{ as } k \to \infty. \tag{5.93}$$

Equation (5.93) is, in effect, a condition on the smoothness of the function f sufficient to ensure that we cannot improve on the accuracy of the first-kind Chebyshev expansion by expanding in ultraspherical polynomials $P_k^{(\alpha)}$ for any $\alpha > -\frac{1}{2}$ (and so, in particular, in Legendre polynomials or in second-kind Chebyshev polynomials). Light's analysis, however, still does not exclude the possibility that we could get faster convergence to such a function f by taking $0 < \alpha < -\frac{1}{2}$, although we do not believe that anyone has yet constructed a function f for which this is the case.

5.7 The error of a truncated Chebyshev expansion

There are many applications of Chebyshev polynomials, especially to ordinary and partial differential equations, where we are approximating a function that is continuously differentiable, finitely or infinitely many times. If this is the case, then Chebyshev expansion converges very rapidly, as the following theorems show.

Theorem 5.14 *If the function $f(x)$ has $m+1$ continuous derivatives on $[-1,1]$, then $\left|f(x) - S_n^T f(x)\right| = O(n^{-m})$ for all x in $[-1,1]$.*

We can prove this using *Peano's theorem* (Davis 1961, p.70) as a lemma.

Lemma 5.15 (Peano, 1913) *Let L be a bounded linear functional on the space $\mathcal{C}^{m+1}[a,b]$ of functions with $m+1$ continuous derivatives, such that $Lp_m = 0$ for every polynomial p_m in Π_m. Then, for all $f \in \mathcal{C}^{m+1}[a,b]$,*

$$Lf = \int_a^b f^{(m+1)}(t) K(t)\,\mathrm{d}t \tag{5.94}$$

where

$$K(t) = \frac{1}{m!} L(\cdot - t)_+^m. \tag{5.95}$$

Here the notation $(\cdot)_+^m$ means

$$(x-t)_+^m := \begin{cases} (x-t)^m, & x \geq t \\ 0, & x < t. \end{cases} \tag{5.96}$$

Proof: (of Theorem 5.14)

Let $f \in \mathcal{C}^{m+1}[-1,1]$. If $S_n^T f$, as in (5.67), is the Chebyshev partial sum of degree $n \geq m$ of f, then the operator L_n, defined for any fixed value $x \in [-1,1]$ by the relationship

$$L_n f := (S_n^T f)(x) - f(x), \tag{5.97}$$

is a bounded linear functional on $\mathcal{C}^{m+1}[-1,1]$. Since $S_n^T p_m \equiv p_m$ for every polynomial in Π_m, it follows that $L_n p_m = 0$ for every such polynomial. Using Peano's theorem, we deduce that

$$(S_n^T f)(x) - f(x) = \int_{-1}^{1} f^{(m+1)}(t) K_n(x,t)\,\mathrm{d}t \tag{5.98}$$

where

$$K_n(x,t) = \frac{1}{m!}\{S_n^T (x-t)_+^m - (x-t)_+^m\}. \tag{5.99}$$

We note that in (5.99) the operator S_n^T must be regarded as acting on $(x-t)_+^m$ as a function of x, treating t as constant; thus, explicitly, $S_n^T(x-t)_+^m = \sum_{k=0}^{n}{}' c_{km} T_k(x)$ where

$$c_{km} = \frac{2}{\pi}\int_t^1 \frac{(x-t)^m T_k(x)}{\sqrt{1-x^2}}\,\mathrm{d}x \tag{5.100}$$

or, writing $x = \cos\theta$ and $t = \cos\phi$,

$$K_n(\cos\theta, \cos\phi) = \frac{1}{m!}\left\{ \sum_{k=0}^{n}{}' c_{km}\cos k\theta - (\cos\theta - \cos\phi)_+^m \right\} \tag{5.101}$$

where

$$c_{km} = \frac{2}{\pi}\int_0^{\phi} (\cos\theta - \cos\phi)^m \cos k\theta \;\mathrm{d}\theta. \tag{5.102}$$

Now it can be shown that $c_{km} = O(k^{-m-1})$ as $k \to \infty$. It follows that

$$\left| S_n^T(x-t)_+^m - (x-t)_+^m \right| = \left| \sum_{k=n+1}^{\infty} c_{km} T_k(x) \right| \le \sum_{k=n+1}^{\infty} |c_{km}| = O(n^{-m})$$

and hence finally, using (5.99) and (5.98),

$$\left| (S_n^T f)(x) - f(x) \right| = O(n^{-m}).$$

This completes the proof. ●●

If f is infinitely differentiable, clearly convergence is faster than $O(n^{-m})$ however big we take m. In some circumstances we can say even more than this, as the following theorem shows.

Theorem 5.16 *If the function $f(x)$ can be extended to a function $f(z)$ analytic on the ellipse E_r of (1.44), where $r > 1$, then $\left| f(x) - S_n^T f(x) \right| = O(r^{-n})$ for all x in $[-1,1]$.*

Proof: Suppose that

$$M = \sup\{|f(z)| : z \in E_r\}. \tag{5.103}$$

The Chebyshev expansion will converge, so that we can express the error as

$$f(x) - (S_n^T f)(x) = \sum_{k=n+1}^{\infty} \frac{2}{\pi}\int_{-1}^{1} (1-y^2)^{-\frac{1}{2}} f(y) T_k(y) T_k(x)\,\mathrm{d}y. \tag{5.104}$$

Using the conformal mapping of Section 1.4.1, with

$$x = \tfrac{1}{2}(\xi + \xi^{-1}), \quad f(x) = g(\xi) = g(\xi^{-1})$$

(so that $|g(\zeta)| \le M$ for $r^{-1} \le |\zeta| \le r$), and remembering that integration around the unit circle C_1 in the ξ-plane corresponds to integration *twice* along the interval $[-1, 1]$ in the x-plane (in opposite directions, but taking different branches of the square root function), we get

$$\begin{aligned}
&f(x) - (S_n^T f)(x) = \\
&= \sum_{k=n+1}^{\infty} \frac{1}{4\mathrm{i}\pi} \oint_{C_1} g(\eta)(\eta^k + \eta^{-k})(\xi^k + \xi^{-k}) \frac{\mathrm{d}\eta}{\eta} \\
&= \sum_{k=n+1}^{\infty} \frac{1}{4\mathrm{i}\pi} \left[\oint_{C_r} g(\eta)\eta^{-k}(\xi^k + \xi^{-k}) \frac{\mathrm{d}\eta}{\eta} + \oint_{C_{1/r}} g(\eta)\eta^{k}(\xi^k + \xi^{-k}) \frac{\mathrm{d}\eta}{\eta} \right] \\
&\qquad \text{— since all parts of the integrand are analytic between } C_r \text{ and } C_{1/r} \\
&= \sum_{k=n+1}^{\infty} \frac{1}{2\mathrm{i}\pi} \oint_{C_r} g(\eta)\eta^{-k}(\xi^k + \xi^{-k}) \frac{\mathrm{d}\eta}{\eta} \\
&\qquad \text{— replacing } \eta \text{ by } \eta^{-1} \text{ in the second integral, and using } g(\eta) = g(\eta^{-1}) \\
&= \frac{1}{2\mathrm{i}\pi} \oint_{C_r} g(\eta) \left(\frac{\xi^{n+1}\eta^{-n-1}}{1 - \xi\eta^{-1}} + \frac{\xi^{-n-1}\eta^{-n-1}}{1 - \xi^{-1}\eta^{-1}} \right)) \frac{\mathrm{d}\eta}{\eta},
\end{aligned} \tag{5.105}$$

where $|\xi| = 1$ when $x \in [-1, 1]$. Therefore

$$\left| f(x) - (S_n^T f)(x) \right| \le \frac{M}{r^n(r - 1)}. \tag{5.106}$$

The Chebyshev series therefore converges pointwise at least as fast as r^{-n}. ●●

5.8 Series of second-, third- and fourth-kind polynomials

Clearly we may also form series from Chebyshev polynomials of the other three kinds, and we would then expect to obtain results analogous to those for polynomials of the first kind and, in an appropriate context, further near-best approximations. First, however, we must consider the formation of the series expansions themselves.

5.8.1 Series of second-kind polynomials

A series in $\{U_i(x)\}$ can be found directly by using orthogonality as given by (5.1)–(5.4). If we define a formal expansion of $f(x)$ as

$$f(x) \sim \sum_{i=0}^{\infty} c_i^U U_i(x), \tag{5.107}$$

then

$$c_i^U = \int_{-1}^{1} (1-x^2)^{\frac{1}{2}} f(x) U_i(x)\, \mathrm{d}x \big/ \langle U_i\, ,\, U_i \rangle$$

where

$$\begin{aligned}\langle U_i\, ,\, U_i \rangle &= \int_{-1}^{1} (1-x^2)^{\frac{1}{2}} U_i(x)^2\, \mathrm{d}x \\ &= \int_0^{\pi} \sin^2(i+1)\theta\, \mathrm{d}\theta \\ &= \tfrac{1}{2}\pi.\end{aligned}$$

Thus

$$c_i^U = \frac{1}{\pi}\int_{-1}^{1} (1-x^2)^{\frac{1}{2}} f(x) U_i(x)\, \mathrm{d}x \tag{5.108a}$$

$$= \frac{2}{\pi}\int_0^{\pi} \sin\theta\, \sin(i+1)\theta\, f(\cos\theta)\, \mathrm{d}\theta. \tag{5.108b}$$

For any given $f(x)$, one of these integrals may be computed analytically or (failing that) numerically, for each i, and hence the expansion (5.107) may be constructed.

It is worth noting that from (5.108b) we can get the expression

$$\begin{aligned}c_i^U &= \frac{1}{\pi}\int_0^{\pi} \{\cos i\theta - \cos(i+2)\theta\} f(\cos\theta)\, \mathrm{d}\theta \\ &= \tfrac{1}{2}\{c_i^T - c_{i+2}^T\},\end{aligned} \tag{5.109}$$

where $\{c_i^T\}$ are the coefficients of the first-kind Chebyshev series (5.6) of $f(x)$. This conclusion could equally well have been deduced from the relationship (1.7)

$$U_n(x) - U_{n-2}(x) = 2T_n(x).$$

Thus a second-kind expansion can be derived directly from a first-kind expansion (but not *vice versa*).

Another way of obtaining a second-kind expansion may be by differentiating a first-kind expansion, using the relation (2.33)

$$T_n'(x) = nU_{n-1}(x).$$

For example, the expansion (5.18), for $z = 1$,

$$\mathrm{e}^x \sim I_0(1) + 2\sum_{i=1}^{\infty} I_i(1) T_i(x)$$

immediately yields on differentiation

$$\mathrm{e}^x \sim 2\sum_{i=0}^{\infty}(i+1)I_{i+1}(1)U_i(x), \tag{5.110}$$

where I_i is the modified Bessel function.

(Note that we have $\sum$ and not $\sum'$ in (5.110) — that is, the U_0 coefficient is *not* halved in the summation. It is only in sums of first-kind polynomials that this halving is naturally required.)

Operating in reverse, we may generate a first-kind expansion by integrating a given second-kind expansion. In fact, this is a good approach to the indefinite integration of a given function, since it yields a first-kind expansion of the integral and hence its partial sums are good approximations in the L_∞ sense. We shall discuss this in more depth later.

It can also sometimes be advantageous to weight a second-kind expansion by $\sqrt{1-x^2}$. For example, the expansion

$$\sqrt{1-x^2}\,f(x) \sim \sum_{i=0}^{\infty} c_i^U \sqrt{1-x^2}\,U_i(x), \tag{5.111}$$

where c_i^U are defined by (5.108a) or (5.108b), can be expected to have good convergence properties provided that $f(x)$ is suitably smooth, since each term in the expansion has a minimax property among polynomials weighted by $\sqrt{1-x^2}$.

5.8.2 Series of third-kind polynomials

A function may also be directly expanded in third-kind polynomials in the form

$$f(x) \sim \sum_{i=0}^{\infty} c_i^V V_i(x). \tag{5.112}$$

Now if $x = \cos\theta$ then

$$V_i(x) = \frac{\cos(i+\frac{1}{2})\theta}{\cos\frac{1}{2}\theta}$$

and

$$\mathrm{d}x = -\sin\theta\,\mathrm{d}\theta = 2\sin\tfrac{1}{2}\theta\ \cos\tfrac{1}{2}\theta\,\mathrm{d}\theta.$$

Hence

$$\begin{aligned} c_i^V &= \frac{\int_{-1}^{1}(1+x)^{\frac{1}{2}}(1-x)^{-\frac{1}{2}}V_i(x)f(x)\,\mathrm{d}x}{\int_{-1}^{1}(1+x)^{\frac{1}{2}}(1-x)^{-\frac{1}{2}}V_i(x)^2\,\mathrm{d}x} \\ &= \frac{\int_0^\pi 2\cos\frac{1}{2}\theta\ \cos(i+\frac{1}{2})\theta\ f(\cos\theta)\,\mathrm{d}\theta}{\int_0^\pi 2\cos\frac{1}{2}\theta\ \cos^2(i+\frac{1}{2})\theta\,\mathrm{d}\theta}. \end{aligned}$$

Thus

$$c_i^V = \frac{1}{\pi}\int_0^\pi \{\cos i\theta + \cos(i+1)\theta\} f(\cos\theta)\,\mathrm{d}\theta = \tfrac{1}{2}\{c_i^T + c_{i+1}^T\} \tag{5.113}$$

(which is consistent with (1.20)); the expansion coefficients may hence be calculated either directly or indirectly.

For example, suppose

$$f(x) = 2^{-\frac{1}{2}}(1-x)^{\frac{1}{2}},$$

so that $f(\cos\theta) = \sin\frac{1}{2}\theta$. Then

$$\begin{aligned} c_i^V &= \frac{1}{\pi}\int_0^\pi \sin\theta\ \cos(i+\tfrac{1}{2})\theta\,\mathrm{d}\theta \\ &= \frac{1}{\pi}\int_0^{\frac{1}{2}\pi} 2\sin 2\phi\ \cos(2i+1)\phi\,\mathrm{d}\phi \\ &= \frac{1}{\pi}\int_0^{\frac{1}{2}\pi} [\sin(2i+3)\phi - \sin(2i-1)\phi]\,\mathrm{d}\phi. \end{aligned}$$

Thus

$$c_i^V = -\frac{1}{\pi}\left(\frac{1}{2i-1} - \frac{1}{2i+3}\right) = -\frac{4}{\pi}\frac{1}{(2i-1)(2i+3)}, \tag{5.114}$$

and we obtain the expansion

$$2^{-\frac{1}{2}}(1-x)^{\frac{1}{2}} \sim -\frac{4}{\pi}\sum_{i=0}^{\infty}\frac{1}{(2i-1)(2i+3)}V_i(x). \tag{5.115}$$

In fact, any third-kind expansion such as (5.115) can be directly related to a first-kind expansion in polynomials of odd degree, as follows. Write $x = 2u^2 - 1$, so that $u = \cos\frac{1}{2}\theta$. We observe that, since (1.15) holds, namely

$$V_n(x) = u^{-1}T_{2n+1}(u),$$

the third-kind expansion (5.112) gives

$$uf(2u^2-1) \sim \sum_{i=0}^{\infty} c_i^V T_{2i+1}(u). \tag{5.116}$$

Thus, since the function $f(2u^2-1)$ is an even function of u, so that the left-hand side of (5.116) is odd, the right-hand side must be the first-kind Chebyshev expansion of $uf(2u^2-1)$, all of whose even-order terms must vanish.

Indeed, for the specific example

$$f(x) = 2^{-\frac{1}{2}}(1-x)^{\frac{1}{2}},$$

we have

$$uf(2u^2-1) = u\sqrt{1-u^2}$$

and hence we obtain the expansion

$$x\sqrt{1-x^2} \sim \sum_{j=0}^{\infty} c_i^V T_{2i+1}(x) \tag{5.117}$$

where c_i^V is given by (5.114).

Fourth-kind expansions may be obtained in a similar way to third-kind expansions, simply by reversing the sign of x.

5.8.3 Multivariate Chebyshev series

All the near-minimax results for first-, second-, third- and fourth-kind polynomials extend to multivariate functions on hypercubes, with the Lebesgue constant becoming a product of the component univariate Lebesgue constants—see Mason (1980, 1982) for details.

5.9 Lacunary Chebyshev series

A particularly interesting, if somewhat academic, type of Chebyshev series is a *'lacunary'* series, in which non-zero terms occur progressively less often as the series develops. For example, the series

$$\begin{aligned} f(x) &= T_0(x) + 0.1\,T_1(x) + 0.01\,T_3(x) + 0.001\,T_9(x) + 0.0001\,T_{27}(x) + \cdots \\ &= T_0(x) + \sum_{k=0}^{\infty} (0.1)^{k+1} T_{3^k}(x) \end{aligned} \tag{5.118}$$

is such a series, since the degrees of the Chebyshev polynomials that occur grow as powers of 3. This particular series is also uniformly convergent, being absolutely bounded by the geometric progression

$$\sum_{k=0}^{\infty} (0.1)^k = \tfrac{10}{9}.$$

The series (5.118) has the remarkable property that its partial sum of degree $N = 3^n$, namely

$$p_N := T_0(x) + \sum_{k=0}^{n} (0.1)^{k+1} T_{3^k}(x), \tag{5.119}$$

is a minimax approximation of degree $(3^{n+1} - 1)$ to $f(x)$, since the error of this approximation is

$$e_N = f(x) - p_N(x) = \sum_{k=n+1}^{\infty} (0.1)^{k+1} T_{3^k}(x),$$

and the equioscillating extrema of each of the polynomials $T_{3^k}(x)$ for $k > n+1$ include $3^{n+1} + 1$ extrema that coincide in position and sign with those of $T_{3^{n+1}}(x)$; therefore their sum has equioscillating extrema at these same points, and we can apply the alternation theorem (Theorem 3.4).

Generalising the above result, we can prove the following lemma and theorem.

Lemma 5.17 *If r is an odd integer greater than 2, the polynomials $T_{r^k}(x)$, $(k = n, n+1, \ldots)$ have a common set of $r^n + 1$ extrema of equal (unit) magnitude and the same alternating signs at the points $x = \cos k\pi/r^n$, $(k = 0, 1, \ldots, r^n)$.*

Theorem 5.18 *If r is an odd integer greater than 2, and $\sum_{k=0}^{\infty} |a_k|$ is convergent, then the minimax polynomial approximation of every degree between r^n and $r^{n+1} - 1$ inclusive to the continuous function*

$$f(x) = \sum_{k=0}^{\infty} a_k T_{r^k}(x) \tag{5.120}$$

is given by the partial sum of degree r^n of (5.120).

A similar result to Theorem 5.18, for $\mathcal{L}_1$ approximation by a lacunary series in $U_{r^k-1}(x)$ subject to restrictions on r and a_k, based on Theorem 6.10 below, is given by Freilich & Mason (1971) and Mason (1984).

5.10 Chebyshev series in the complex domain

If the function $f(z)$ is analytic within and on the elliptic contour E_r (4.81) in the complex plane, which surrounds the real interval $[-1, 1]$ and has the points $z = \pm 1$ as its foci, then we may define alternative orthogonal expansions in Chebyshev polynomials, using the inner product (4.83)

$$\langle f\,,\, g\rangle := \oint_{E_r} f(z)\overline{g(z)}\, |\mu(z)|\, |\mathrm{d}z|\,, \tag{5.121}$$

of Section 4.9 in place of (5.1).

Specifically, in the case of polynomials of the first kind, we can construct the expansion

$$f(z) \sim \sum_{k=0}^{\infty}{}' c_k T_k(z) \tag{5.122}$$

where (taking the value of the denominator from (4.85a))

$$c_k = \frac{\langle f\,,\, T_k\rangle}{\langle T_k\,,\, T_k\rangle} = \frac{2}{\pi(r^{2k} + r^{-2k})} \oint_{E_r} f(z)\overline{T_k(z)} \left| \frac{\mathrm{d}z}{\sqrt{1-z^2}} \right| . \tag{5.123}$$

As in Section 4.9, we make the substitution (4.75)

$$z = \tfrac{1}{2}(w + w^{-1}), \tag{5.124}$$

under which the ellipse E_r in the z-plane is the image of the circle C_r of radius $r > 1$ in the w-plane:

$$C_r = \{w : w = r\mathrm{e}^{\mathrm{i}\theta}, \theta \text{ real}\}.$$

Then (4.77)

$$T_k(z) = \tfrac{1}{2}(w^k + w^{-k})$$

and hence, for w on C_r,

$$\overline{T_k(z)} = \tfrac{1}{2}(\overline{w}^k + \overline{w}^{-k}) = \tfrac{1}{2}(r^{2k}w^{-k} + r^{-2k}w^k). \tag{5.125}$$

For w on C_r, we also have

$$\left| \frac{\mathrm{d}z}{\sqrt{1-z^2}} \right| = \mathrm{d}\theta = \frac{\mathrm{d}w}{\mathrm{i}w}. \tag{5.126}$$

Define the function g such that for all w

$$g(w) = f(z) \equiv f(\tfrac{1}{2}(w + w^{-1})); \tag{5.127}$$

then we note that $g(w)$ will be analytic in the annulus between the circles C_r and $C_{r^{-1}}$, and that we must have $g(w^{-1}) = g(w)$.

Now we have

$$\begin{aligned} c_k &= \frac{2}{\pi(r^{2k} + r^{-2k})} \oint_{E_r} f(z)\overline{T_k(z)} \left| \frac{\mathrm{d}z}{\sqrt{1-z^2}} \right| \\ &= \frac{1}{\pi(r^{2k} + r^{-2k})} \oint_{C_r} g(w)(r^{2k}w^{-k} + r^{-2k}w^k) \frac{\mathrm{d}w}{\mathrm{i}w}. \end{aligned} \tag{5.128}$$

Since the function $g(w)$ is analytic in the annulus between the circles C_r and $C_{r^{-1}}$, and satisfies $g(w^{-1}) = g(w)$, we can show, by applying Cauchy's theorem to this annulus and then changing variable from w to w^{-1}, that

$$\oint_{C_r} g(w)w^k \frac{\mathrm{d}w}{\mathrm{i}w} = \oint_{C_{r^{-1}}} g(w)w^k \frac{\mathrm{d}w}{\mathrm{i}w} = \\ = \oint_{C_r} g(w^{-1})w^{-k} \frac{\mathrm{d}w}{\mathrm{i}w} = \oint_{C_r} g(w)w^{-k} \frac{\mathrm{d}w}{\mathrm{i}w}. \quad (5.129)$$

Combining (5.128) and (5.129), we get

$$c_k = \frac{1}{\mathrm{i}\pi} \oint_{C_r} g(w)w^k \frac{\mathrm{d}w}{w}. \quad (5.130)$$

The expansion (5.122) thus becomes

$$f(z) \sim {\sum_{k=0}^{\infty}}' \left\{ \frac{1}{\mathrm{i}\pi} \oint_{C_r} g(w)w^k \frac{\mathrm{d}w}{w} \right\} T_k(z) \quad (5.131)$$

or

$$g(\zeta) \sim {\sum_{k=0}^{\infty}}' \left\{ \frac{1}{2\mathrm{i}\pi} \oint_{C_r} g(w)w^k \frac{\mathrm{d}w}{w} \right\} (\zeta^k + \zeta^{-k}) \\ = \sum_{k=-\infty}^{\infty} \frac{1}{2\mathrm{i}\pi} \left\{ \oint_{C_r} g(w)w^k \frac{\mathrm{d}w}{w} \right\} \zeta^k, \quad (5.132)$$

making use of (5.129) again.

We may now observe that (5.132) is just the Laurent expansion of $g(\zeta)$ in positive and negative powers of ζ. So, just as in the real case we were able to identify the Chebyshev series of the first kind with a Fourier series, in the complex case we can identify it with a Laurent series.

5.10.1 Chebyshev–Padé approximations

There is a huge literature on Padé approximants (Padé 1892)—rational functions whose power series expansions agree with those of a given function to as many terms as possible—mainly because these approximants often converge in regions beyond the radius of convergence of the power series. Comparatively little has been written (Gragg 1977, Chisholm & Common 1980, Trefethen & Gutknecht 1987, for a few examples) on analogous approximations by ratios of sums of Chebyshev polynomials. However, the Chebyshev–Padé approximant seems closely related to the traditional Padé table (Gragg & Johnson 1974), because it is most easily derived from the link to Laurent series via the property

$$T_n(z) = \tfrac{1}{2}(z^n + z^{-n}),$$

w and z being related by (5.124), so that we may match

$$\frac{\sum_{k=0}^{p}{}' a_k \tfrac{1}{2}(w^k + w^{-k})}{\sum_{k=0}^{q}{}' b_k \tfrac{1}{2}(w^k + w^{-k})} \quad \text{and} \quad \sum_{k=0}^{\infty}{}' c_k \tfrac{1}{2}(w^k + w^{-k})$$

up to the term in $w^{p+q+1} + w^{-(p+q+1)}$, by multiplying through by the denominator and equating the coefficients of positive (or, equivalently, negative) and zero powers of w.

There has also been work on derivations expressed entirely in terms of Chebyshev polynomials; the first that we are aware of is that of Maehly (1960) and a more efficient procedure, based on only $p + q + 1$ values of c_k, is given by Clenshaw & Lord (1974).

5.11 Problems for Chapter 5

1. Verify the Chebyshev expansions of $\operatorname{sgn} x$, $|x|$ and $\delta(x)$ quoted in (5.11), (5.12) and (5.24).

2. If $\hat{c}_i$ denotes the trapezium-rule approximation to c_i defined by the right-hand side of (5.38), x_k being taken at the zeros of $T_n(x)$, show that

$$\begin{aligned} \hat{c}_n &= 0, \\ \hat{c}_{2n\pm i} &= -\hat{c}_i, \\ \hat{c}_{4n-i} &= \hat{c}_i. \end{aligned}$$

3. Show that the mapping S_n^T, defined so that $S_n^T f$ is the nth partial sum of the Chebyshev series expansion of f, is a projection.

4. Prove (5.50):

   (a) directly;

   (b) by applying (1.14) and (1.15) to Exercise (3a) of Chapter 2 to deduce that
   $$\sum_{k=0}^{n}{}' T_k(x) = \tfrac{1}{2} W_n(x)$$
   and then making the substitution $x = \cos s$.

5. If λ_n is given by (5.71) show, using the inequality $\left|\sin \tfrac{1}{2}t\right| \le \left|\tfrac{1}{2}t\right|$, that $\lambda_n > \dfrac{4}{\pi^2} \log n$.

6. With g_D as defined by (5.82), show that if τ is sufficiently small then the function g_C defined by

$$g_C(t) := \frac{1}{2\tau} \int_{t-\tau}^{t+\tau} g_D(s)\,\mathrm{d}s$$

has all the properties required to complete the proof in Section 5.5.1, namely that g_C is continuous, even and periodic, $\|g_C\|_\infty \le 1 + \epsilon$ and $\|g_C - g_D\|_1 < \epsilon$.

7. Assuming that $f(z)$ is real when z is real, show that the coefficients c_k defined by (5.123) are the same as those defined by (5.7).

8. Consider the partial sum of degree n of the first kind Chebyshev series expansion of a function $f(z)$, analytic on the interior of the ellipse E_r : $\left|z + \sqrt{z^2 - 1}\right| = r$ $(r > 1)$ and continuous on E_r. Show that this sum maps under $z = \frac{1}{2}(w + w^{-1})$ into the partial sum of an even Laurent series expansion of the form $\frac{1}{2}\sum_{-n}^{n} c_k w^k$, where $c_{-k} = c_k$.

9. Obtain Cauchy's integral formula for the coefficients c_k and Dirichlet's formula for the partial sum of the Laurent series, and interpret your results for a Chebyshev series.

10. Following the lines of argument of Problems 8 and 9 above, derive partial sums of second kind Chebyshev series expansions of $(z^2 - 1)^{\frac{1}{2}} f(z)$ and a related odd Laurent series expansion with $c_{-k} = -c_k$. Again determine integral formulae for the coefficients and partial sums.

11. Using the Dirichlet formula of Problem 9, either for the Chebyshev series or for the related Laurent series, show that the partial sum is near-minimax on E_r within a relative distance λ_n.

12. Supposing that

$$G(x) = g_1(x) + \sqrt{1 - x^2}\, g_2(x) + \sqrt{\frac{1+x}{2}}\, g_3(x) + \sqrt{\frac{1-x}{2}}\, g_4(x),$$

where g_1, g_2, g_3, g_4 are continuously differentiable, and that

$$g_1(x) \sim \sideset{}{'}\sum_{r=0}^{n} a_{2r} T_r(x) + \cdots, \quad g_2(x) \sim \sum_{r=1}^{n} b_{2r} U_{r-1}(x) + \cdots,$$

$$g_3(x) \sim \sum_{r=0}^{n-1} a_{2r+1} V_r(x) + \cdots, \quad g_4(x) \sim \sum_{r=0}^{n-1} b_{2r+1} W_r(x) + \cdots,$$

determine the form of $F(\theta) = G(\cos\theta)$. Deduce that

$$F(2\theta) = \sideset{}{'}\sum_{k=0}^{2n} (a_k \cos k\theta + b_k \sin k\theta) + \cdots.$$

Discuss the implications of this result in terms of separating a function into four component singular functions, each expanded in a different kind of Chebyshev series.